Ten Years after the Financial Crisis:
What Have We Learned from the Renaissance in Fiscal Research?*

By
Valerie A. Ramey
University of California, San Diego and NBER

September 15, 2018

Abstract

This paper takes stock of what we have learned from the “Renaissance” in fiscal research in the ten years since the financial crisis. I first summarize the new innovations in methodology and discuss the various strengths and weaknesses of the main approaches. Reviewing the estimates, I come to the surprising conclusion that the bulk of the estimates for average spending and tax change multipliers lie in a fairly narrow range, 0.6 to 1 for spending multipliers and -2 to -3 for tax change multipliers. However, I identify economic circumstances in which multipliers lie outside those ranges. I conclude by reviewing the debate on whether multipliers were higher on the stimulus spending in the U.S. and the fiscal consolidations in Europe.

* I am grateful for helpful comments from Alberto Alesina, Gabriel Chodorow-Reich, Martin Eichenbaum, Carlo Favero, Francesco Giavazzi, Karel Mertens, Maury Obstfeld, Linda Tesar, Sarah Zubairy, and participants at the July 2018 NBER Conference “Global Financial Crisis @10.”
When the financial crisis hit ten years ago and interest rates fell to their lower bound, policymakers around the world turned to fiscal stimulus packages in order to prevent their economies from freefalling into another Great Depression. But as declining GDP and tax revenues led to deteriorating government budgets and worries about sovereign debt, numerous countries abandoned stimulus packages and instead adopted fiscal consolidation measures. While attempting to forecast the impacts of these various fiscal programs, policymakers and academics were surprised to learn that not only was there no consensus about the size of the effects of fiscal policy but that there had not even been much research on the topic since the 1960s. This situation changed quickly as armies of researchers across many countries turned their attention to this important, but long-neglected, topic. Indeed, a positive by-product of the financial crisis has been a renaissance in research on the macroeconomic effects of fiscal policy.

In the last ten years, important progress has been made on all three important methodological fronts: theory, empirical methods, and data. On the theory front, we now have a much better understanding of how deviations from the classic Baxter-King (1993) neoclassical benchmark affect predicted multipliers. The theoretical innovations include the analysis of the effects of sticky prices, hand-to-mouth consumers, lower bounds on policy interest rates, currency unions, the type of financing, and anticipations on the reactions of macroeconomic variables to fiscal policy. The contributions in the realm of empirical methods include important new ways to identify exogenous variation in policy, standardization of methods for computing multipliers, and the incorporation of state dependence. On the data front, researchers now have much more data that can be used to estimate effects. In addition to newly constructed historical and cross-sectional data sets, researchers are also exploiting the rich new data created by the variety of policymakers’ fiscal responses to the crisis. All of these advancements offer the
potential to estimate the effects of government spending with more precision and with a better understanding of how the effects depend on the particular context.

This paper takes a snapshot of the state of knowledge about the effects of fiscal policy ten years after the global financial crisis. In 2011, I surveyed the pre-crisis and early crisis literature in the *Journal of Economic Literature*. In that paper, which focused only on temporary, deficit-financed increases in government purchases, I concluded based on the evidence available from U.S. data at that time that the multiplier was probably between 0.8 to 1.5, but that reasonable people could argue that the data did not reject a range from 0.5 to 2. The current paper refines those estimates and broadens the inquiry to consider the effects of tax and transfer policy, as well as the effects of fiscal consolidations, in developed countries. However, attention is still limited to the short- or medium run effects since the methods for estimating long-run effects are quite different.

My summary of the current state of knowledge about the effects of fiscal policies is as follows.

**Government purchases multipliers**

- On average, multipliers on general government purchases in developed countries are positive but less than or equal to unity, meaning that government purchases raise GDP but do not stimulate additional private activity and may actually crowd it out. The bulk of the estimates across the leading methods of estimation and samples lie in a surprisingly narrow range of 0.6 to 1. However, this range widens once one distinguishes country characteristics, such as the exchange rate regime, and the type of government spending, such as infrastructure spending.
• The evidence for higher spending multipliers during recessions or times of high unemployment is fragile. While some time series and panel estimates suggest multipliers above one, the most robust results suggest multipliers of one or below during these periods. Moreover, standard New Keynesian quantitative models do not predict procyclical multipliers.

• The evidence for higher spending multipliers during periods in which monetary policy is very accommodative, such as zero lower bound periods, is somewhat stronger. Recent time series estimates for the U.S. and Japan suggest that multipliers could be 1.5 or higher during those times. Estimated and calibrated New Keynesian models for the U.S. and Europe also imply higher multipliers under certain conditions.

Tax rate change multipliers

• On average, narrative methods for tax rate changes yield multiplier estimates that are surprisingly large and surprisingly uniform across a number of countries. The bulk of the empirical estimates vary between -2 and -3. However, most calibrated and estimated DSGE models imply smaller multipliers, typically below unity for both labor and capital tax multipliers. Thus, there is currently a gap between the estimates implied by the two leading methods.

• Time series evidence, theory, and estimated New Keynesian DSGE models all point tax multipliers being greater in magnitude during expansions than in recessions. Thus, tax multipliers appear to be procyclical if anything.

Multipliers in the Wake of the Financial Crisis
• The evidence for larger national multipliers on the Obama stimulus package is at best weak. Quantitative New Keynesian models do not find larger multipliers. The larger multipliers estimated on cross-state data shrink once one adjusts them to make them nationally representative.

• The latest studies on multipliers during the fiscal consolidations in Europe suggest that they were not higher than usual.

The outline of the rest of the paper is as follows. It begins by briefly reviewing how theory highlights the dependence of the size of the multipliers on numerous features of the policy and the economy. The next section summarizes strengths and weaknesses of each of the leading empirical approaches to identifying exogenous shifts in fiscal policy. The paper then goes on to highlight the innovations of the last ten years in estimating fiscal multipliers. One of the interesting findings is that the wide range of multipliers reported earlier is reduced significantly once methods for calculating multipliers are standardized. The next section reviews the leading estimates of spending and tax multipliers, including those from aggregate time series, estimated theoretical models, and from subnational units and households. It also discusses the complexities of drawing aggregate inferences from parameters estimated on household data. The penultimate section asks what we know about whether multipliers were higher in the wake of the financial crisis and the final section concludes.

What Does Theory Predict?

If we simply want to know how much GDP changes if we increase government spending by one dollar or reduce tax rates by one percentage point, why do we need theory? As this
section explains, the description of the policy in the previous sentence is incomplete. Theory tells us that there is not one government spending or tax multiplier. Rather, the impact on output and other variables potentially depends on (i) the persistence of the change, (ii) the type of spending or taxes that changed, (iii) how the policy was financed, (iv) whether it was anticipated, (v) how the policy was distributed across potentially heterogeneous agents, (vi) how monetary policy reacted, (vii) the state of the economy when the policy took effect, and (viii) what other features characterize the economy in question, such as the level of development, the exchange rate regime, and the openness. Because policymakers cannot conduct randomized control trials, virtually all estimates are based on time series, narrative, or natural experiment identification using convenience samples determined by historical happenstance. In order to understand whether a particular estimate of fiscal effects is suitable for use in predicting the impact of a proposed policy, one must understand how the current circumstances differ from those present in the sample that was used to generate the estimate.

Most researchers and policymakers’ first exposure to the theoretical effects of fiscal policy is the Keynesian-Cross closed economy model of undergraduate textbooks. In this simple model, which assumes that GDP is demand-determined, the government spending multiplier is the inverse of one minus the marginal propensity to consume. A marginal propensity to consume of 0.5 yields a multiplier of two. Since taxes enter only through disposable income, the tax multipliers are smaller than the spending multipliers. Expansion of the model to consider the marginal propensity to import, tax rates, and monetary policy reduces those simple multipliers.

Neoclassical models with variable labor supply and capital stock also predict positive spending multipliers and negative (distortionary) tax multipliers, but the mechanism is completely different from the one at the heart of the traditional Keynesian model. An increase in
government spending has a negative wealth effect since the government is extracting resources from the private sector. This negative wealth effect raises GDP because it causes households to supply more labor. Distortionary tax rate changes can have potentially large effects. Contrary to the simple Keynesian model, they work through “supply side” channels.

The New Keynesian dynamic stochastic general equilibrium (DSGE) models meld the insights from the traditional Keynesian and neoclassical approaches in a rigorous way. The standard representative agent sticky price New Keynesian model in which monetary policy follows a standard Taylor rule tends to produce multipliers below one for government spending. Models that add sticky wages and workers who are “off their labor supply curves” generate larger multipliers. In the last decade, the representative agent models have been expanded to include heterogeneous agents. In these models, a share of the consumers behaves like the traditional Keynesian consumers rather than being permanent income consumers. These agents consume according to “rule-of-thumb” or “hand-to-mouth” rules, meaning they immediately consume any income they receive, whether it is temporary or permanent. Alternatively, the models have explored the effects of fiscal policy when monetary policy deviates from the standard Taylor rule because of interest rates are constrained by the zero lower bound. Both of these extensions result in higher multipliers, often above unity.

All of these models show that the particular multipliers also depend crucially on the other aspects listed above. The persistence of a path of government spending and how it is financed is crucial, as are many other characteristics such as the exchange rate regime. These results must be kept in mind when one is trying to forecast the effects of a specific policy.

A Summary of the Leading Empirical Approaches
Numerous empirical approaches have been used to estimate the effects of fiscal policies. I group these approaches into three broad categories: (1) aggregate country-level time series or panel estimates; (2) estimated or calibrated dynamic stochastic general equilibrium (DSGE) models; and (3) subnational geographic cross-section or panel estimates.

The first two categories, time series evidence at the national level and estimated/calibrated DSGE models, share the advantage that the estimates produced are directly informative about the national-level multipliers that are the focus of most policymakers. The time series approach has the advantage of not being tied to a particular structural model, as required by the DSGE model approach. On the other hand, the DSGE model approach can be used to perform counterfactuals since it estimates structural parameters.

These two approaches share some of the same weaknesses, though. Identification of macroeconomic parameters is always difficult and the estimation of the aggregate effects of fiscal policy is no exception. The time series approach requires exogenous variation in policy. The leading approaches to identifying this exogenous variation are structural VARS (SVARs) and natural experiment methods combined with narrative methods that use historical documents to create new data series of exogenous changes. Too often, though, the variations that are exogenous yield instruments that are not very relevant, that is, they have low correlation with the fiscal variable they are trying to explain, and the variations that are relevant are not always exogenous or are anticipated in advance.

Although many papers using estimated DSGE models never mention the word identification, identification is as crucial to this approach as it is to any other approach seeking to estimate a causal relationship. The DSGE approach identifies the effects of fiscal policy by using strong assumptions about the theoretical model structure and the time series processes
driving the unobserved shocks. Unfortunately, estimated DSGE models are not immune to weak identification, as pointed out by Fabio Canova and Luca Sala (2009).

The third approach of estimating across subnational units, such as states or provinces, is more similar to applied microeconomics approaches. These approaches typically use the natural experiment approach or Bartik-style instruments that are widely used in applied microeconomics contexts. Similar to the microeconomic context, these analyses at lower levels of aggregation tend to have much stronger identification: the necessary identifying assumptions are typically more plausible and the instruments are relevant. Moreover, these approaches can be used on a variety of datasets. This approach suffers, however, from a key weakness: the estimates produced are not macroeconomic estimates. Why? Any cross-sectional estimating equation includes a constant term, which means that the macroeconomic effects have been netted out. This means that the parameters estimated are only relative effects, i.e., they answer the question: if State A is awarded $1 more in defense prime contracts than the average state, by how much does its employment change relative to the average state? In order to infer the implied national-level effects from the microeconomic estimates, researchers must turn to macroeconomic DSGE models, which, as discussed above, incorporate their own additional identifying assumptions. Thus, there is no “applied micro free lunch” for macroeconomists: identification of macroeconomic effects must always depend on macroeconomic identification assumptions.

To summarize, there are several approaches to estimating the effects of fiscal policy. Each has its strengths, but also its share of weaknesses. Moreover, some of the estimates are more appropriate for forecasting the effects of specific policies under certain conditions than others. For these reasons, it is useful to consider the estimates across the different approaches.
Research Innovations and Lessons Learned During the Last Ten Years

Before the financial crisis, only a few isolated researchers studied the macroeconomic effects of fiscal policy and there were few conferences that brought the researchers together. As a result, different researchers chose different estimation methods and different ways of calculating multipliers from those estimates and there was no agreed upon set of best practices. The situation has changed dramatically since the financial crisis, with many conferences devoted to the study of fiscal policies and much more interaction among researchers studying fiscal policy. As a result, the diffusion of knowledge among researchers has been much faster and the literature has progressed at a very fast pace. In this section, I will highlight some of the new innovations and some of the lessons learned from this literature.

Calculating Multipliers in a Dynamic Environment

Some papers discuss the “wide range” of multiplier estimates. What the literature has come to realize is that differences in reported multiplier estimates are often due not so much to differences in identification methods or samples, but to the methods used to construct multiplier from the raw estimates. In fact, what some researchers call “multipliers” have little to do with the multipliers of interest to policymakers. This section begins with some insights gained over the last decade regarding the computation of multipliers. I begin with spending multipliers and then address a further complication involved with tax multipliers.

Fiscal policy has dynamic effects on output and government budgets. A typical fiscal plan will set into motion a path of spending or taxes over time and GDP will respond dynamically to that path. The multiplier must take into account both the multi-year effects of the
fiscal plan on the government budget, in order to count fully the costs, as well as the multi-year effects on GDP, in order to count fully the benefits.

Computation of multipliers was not a focus of the research in the decades before the financial crisis. Indeed, in my work with Shapiro on the effects of government spending, we did not even mention the word “multiplier” (Ramey and Shapiro (1998)). When describing the patterns of the responses of GDP to spending and tax shocks, Blanchard and Perotti (2002) used the word “multiplier,” but the quantities they calculated were not true dynamic multipliers. In particular, Blanchard and Perotti calculated multipliers as the ratio of the output response at a particular horizon, or at its peak, to the impact effect of the shock on government spending. Many subsequent papers adopted their method, despite the fact that it did not take into account the multi-year path of spending or taxes. Mountford and Uhlig (2009) moved the literature forward by introducing the policy-relevant multipliers, calculated as the present discounted value of the integral of the output response divided by the present discounted value of the integral of the government spending response to the shock. In most applications, the undiscounted integral gives nearly identical multipliers because the timing of the government spending and output responses is very similar. These multipliers are often known as present value or cumulative multipliers.

How much do multiplier estimates differ across these various methods of calculating multipliers? It depends importantly on how much government spending rises after the initial impact. I offer an illustration of a situation in which it makes a big difference. I estimate a Blanchard and Perotti (2002) type structural vector autoregression (SVAR) model over the period 1939:1 – 2015:4 using Ramey and Zubairy’s (2018) data set. The SVAR model contains five endogenous macroeconomic variables: government spending, GDP, and federal tax receipts,
all three deflated by the GDP deflator, divided by population, and in logs; and the 3-month Treasury bill interest rate and inflation (measured as the log change in the GDP deflator). Four lags are included in order to model the dynamics. The exogenous shock to government spending is identified using Blanchard and Perotti’s (2002) method, which assumes that any part of government spending not forecasted by lags of any of the other variables included in the model is an exogenous shock to government spending.

Figure 1 shows the estimated impulse responses of the log of the government spending variable and the log of the GDP variable. The shaded area shows the 95-percent confidence bands. As the graph illustrates, a positive shock to government spending leads both government spending and GDP to jump up on impact, but then to continue to rise, peaking after about a year. Note that because the variables are in log form, the impulse responses show elasticities, not the dollar changes required by multipliers, so multipliers cannot be read directly from the graphs. The standard practice until recently had been to use an ad hoc conversion factor. That is, researchers who specified models using logarithms converted the elasticity estimates, i.e. $\frac{d\ln(Y)}{d\ln(G)}$, to multipliers, $\frac{dY}{dG}$, by multiplying the elasticity estimates by the average of the ratio of GDP to total government spending, $Y/G$, over the sample. This ratio is 4.78 for the sample in this illustration. I will critique the use of these conversion factors shortly.

Figure 2 shows the multipliers calculated three different ways. The highest multiplier is given by Blanchard-Perotti’s method, which I will call a quasi-multiplier. It is calculated as the ratio of the impulse response of output at horizon h to the initial jump in government spending at horizon 0 (multiplied by the average $Y/G$). Their method, shown by the dashed line, essentially traces out a renormalized version of the impulse response of output. In this case it yields multipliers that peak at 2.2 at quarter 6. The Mountford and Uhlig (2009) present value
**cumulative multiplier**, shown by the solid line, uses the ratio of the present value of the integral of impulse response of output to the present value of the integral of the impulse response of government spending up to each horizon h (again multiplied by the average Y/G factor).\(^1\) This multiplier varies between 0.7 and 1, depending on the horizon. The simple cumulative version is almost identical.

A second issue, however, is the practice of converting elasticities with the *ad hoc* conversion factor, the average of \(Y/G\) over the sample. Owyang, Ramey, and Zubairy (2013) discovered biases that could arise from this practice. In their historical sample, \(Y/G\) varied significantly, from 2 to 24, with a mean of 8. Sims and Wolff (2018a,b) also discovered that this practice tends to bias multipliers differentially, making them seem much higher during recessions. The intuition is simple: because GDP is cyclical but government spending is not, \(Y_t/G_t\) is procyclical. However, the practice of using a sample average of \(Y/G\) to convert elasticities to multipliers makes the multipliers appear more countercyclical than they really are. Owyang et al. avoided this problem by using the transformations employed by Hall (2010) and Barro and Redlick (2011): both the change in government spending and the change in GDP are divided by lagged GDP. Another transformation that overcomes the problem is Gordon and Krenn’s (2010) transformation, which divides both government spending and GDP by a measure of potential GDP.

To illustrate the effect of moving from a specification in logarithms that requires the *ad hoc* conversion factor to one that does not, I re-estimate the model using Gordon and Krenn’s transformation for government spending, GDP, and taxes, employing Ramey and Zubairy’s (2018) polynomial trend estimate of potential GDP. The general shape of the estimated impulse responses (not shown) is very similar to those from the log specification, but the directly

---

\(^1\) The discounting uses the average 3-month Treasury bill rate over the sample, 3.6 percent on an annual basis.
estimated multipliers are different. The solid line with diamonds in Figure 2 shows the cumulative multiplier estimates based on the impulse responses from this alternative specification. These directly-estimated multipliers, which do not rely on a conversion factor, are lower and range from 0.8 on impact down to 0.6.⁴

Thus, deceptively small changes in the way the multipliers are calculated can make a very big difference. For this application, using Blanchard and Perotti’s quasi-multiplier for government spending on estimated elasticities requiring an *ad hoc* conversion factor produces a multiplier as high as 2.2. That multiplier falls below 0.8 when the fully dynamic Mountford and Uhlig cumulative multiplier is used on estimates based on data using the Gordon-Krenn transformation. These simple changes could have important consequences for the decisions of policymakers. It should be noted, though, that even the cumulative multipliers do not fully reflect the consequences for the government budget. Tax revenues typically rise when GDP rises. If an increase in government spending raises GDP, then we would expect a rise in tax revenues. Thus, even without an exogenous increase in tax rates, we would expect the government budget deficit to rise less than the total amount of government spending.

The same principles apply to the computation of tax multipliers, with one additional complication. There is little feedback from GDP to government spending, but strong feedback from GDP to tax revenue. As a result, the negative effect of a tax cut on tax revenue is tempered by the feedback from the expansionary effect on output. Indeed, Mertens and Ravn (2013) were not able to compute the multiplier for corporate tax cuts because their large positive impact on GDP resulted in no net effect on tax revenues. Because of the presence of these “top of the Laffer curve” effects in some applications, most papers report multipliers using the tax changes

⁴This bias also affects the multipliers I reported in Ramey (2011a). The cumulative multipliers based on the elasticity estimates and conversion factor were 1.2. However, in Ramey (2013), I found evidence that private spending fell, which is inconsistent with a multiplier above 1.
measured as the legislative forecasts of the expected cumulative effect on tax revenues, not accounting for dynamic feedback from any potential induced GDP changes.

**The Importance of Fiscal Foresight**

An important innovation in the fiscal literature in the last decade is the recognition that many changes in government spending and taxes are announced in advance. In Ramey (2011a) I showed the importance of anticipations for estimating the effects of government spending shocks, particularly involving military spending. For example, I showed that the responses of key variables such as consumption could change signs if researchers ignored the fact that many changes in government spending are anticipated by at least several quarters. A number of papers also show that “shocks” identified in standard ways are predicted by professional forecasts of government spending. On the tax front, House and Shapiro (2006) and Mertens and Ravn (2012) demonstrated the importance of distinguishing between changes in taxes implemented soon after legislation and changes in taxes implemented with a lag after legislation or phased in slowly. Both papers showed that while unanticipated tax cuts have expansionary effects on output, phased in tax cuts depress output during the phase in period since firms and consumers delay their activity until tax rates are lower. Leeper, Walker, and Yang (2013) derived the econometric biases that arise when there is this type of fiscal foresight. As a result of this work, most of the literature tries to address anticipation whenever feasible, either by constructing measures of news (from narratives or bond spreads) or by including professional forecasts of government spending to mitigate the problem.

**Improvements in Fiscal Shock Identification**
Any analysis that seeks to measure a causal effect must confront identification issues because of classic simultaneous equations bias. For example, if governments increase spending in response to a recession, then the simple correlation between government spending and GDP will confound the positive causal effect of government spending on GDP with the negative causal effect of GDP on government spending. Initially, the standard macro approach used was a structural vector autoregression (SVAR). In most applications, it was assumed that the exogenous part of government spending was simply the part of government spending not forecasted by lagged values of spending, GDP, and taxes. To identify exogenous movements in taxes, Blanchard and Perotti (2002) brought in external estimates of the tax revenue elasticity to income, which allowed identification of the component of taxes that was not induced by movements in GDP. Several papers have highlighted potential problems with these widely-used methods. The first problem, discussed above, was the realization that failing to account for fiscal foresight could lead to biased estimates. The second was the demonstration that the tax multiplier estimates were very sensitive to the value of the external tax elasticity estimate used (e.g. Mertens and Ravn (2014), Caldara and Kamps (2017)). These concerns led to the development of other identification methods using natural experiments and narrative methods. As a result, the standard SVAR identification approach is no longer the first resort in the fiscal literature.

In fact, long before SVAR methods were used, Hall (1980) and Barro (1981) used natural experiment methods to assess the effects of exogenous increases in government spending. Arguing that defense spending in the U.S. is typically driven by wars rather than the current state of the economy, they used war-induced government spending to estimate causal effects of government spending in U.S. historical data. Ramey and Shapiro (1998) and numerous other
follow-up papers built on that natural experiment insight about wars. However, this method that works quite well for U.S. data does not export well to other countries since most countries either do not have the substantial fluctuations in defense spending experienced by the U.S. or they have large variations that are accompanied by war-related destruction of the capital stock, which leads to confounding effects.

Other examples of recent fiscal research that use natural experiment methods abound. For example, Acconcia, Corsetti, and Simonelli (2014) used the central government response to Mafia infiltration as an exogenous change in government spending in Italian provinces. Many of the analyses of the Obama stimulus allocation of funds across states used natural experiment methods. Johnson, Parker, and Souleles (2006) and Parker, Souleles, Johnson, and McClelland’s (2013) analysis of marginal propensities to spend out of the temporary rebates of 2001 and 2008 exploited the randomized timing of the mailing of checks to households. The application of these methods has shed significant light on the effects of fiscal policy, particularly at the local and household level.

Romer and Romer (2010) pioneered the use of narrative methods to identify tax changes that are exogenous to the state of the economy. For the post-WWII U.S., they read legislative records to identify tax changes that were due either to inherited deficits or beliefs about their ability to promote long-term growth. Their method is easily exported to other countries and it has now become the standard method for assessing the effects of tax changes across a wide range of countries (e.g. Guajardo, Leigh, and Pescatori (2014)). Mertens and Ravn (2012) improved their measure by splitting their series into anticipated and unanticipated tax changes for the effects of fiscal foresight could be addressed. Alesina, Favero, and Giavazzi (forthcoming) has added to the narrative analysis of fiscal consolidations by creating narrative series of fiscal plans.
As they emphasize, most fiscal consolidations involve multi-year plans and those effects should be studied as a whole rather than as independent year-by-year isolated changes.

An additional innovation in the identification of fiscal shocks has been the recognition of the importance of instrument relevance, i.e. whether the proposed instrument is actually correlated with the variable it is supposed to instrument. While early alarms about weak instruments were raised for macro studies by Nelson and Startz (1990) and for microeconomic studies by Bound, Baker, and Jaeger (1995), macroeconomists began to pay attention to the issue only in the last five to ten years. The SVAR methodology hid the fact that the estimation of multipliers was actually an instrumental variables estimation. Ramey (2016) and Ramey and Zubairy (2018) showed that cumulative multipliers could be estimated in a one-step instrumental variables method based on local projects: cumulative GDP up to horizon $h$ is regressed on cumulative government spending up to horizon $h$, using an SVAR shock or a narrative variable as an instrument. However, that recognition highlighted a widespread problem: many of the exogenous measures of fiscal policy are not very relevant instruments, at all or in some subsamples. For example, the military news variable I first introduced in Ramey (2011a) is a weak instrument for the post-1954 period, as are the alternative measures of defense news of Fisher and Peters (2010) and Ben Zeev and Pappa (2017). In contrast, the Blanchard and Perotti shock, by its nature, is a strong instrument, particularly at short horizons.

In sum, research on the effects of fiscal policy has made significant strides in methodology. The literature now exploits many new data sets. It has imported some of the innovations from the applied microeconomics literature, and has extended them in important ways that account for anticipations and dynamics. Moreover, those estimates are now converted to multipliers defined in a way that is relevant for policymakers.
A Summary of Estimates of Spending and Tax Multipliers

This section summarizes fiscal multiplier estimates obtained from the leading methods. I begin with estimates based on aggregate data. I first review the estimated multipliers on government purchases, initially averages and then by state-dependence, and then move on to the effects of tax changes and transfer payments. I then discuss estimates of the effects of the ARRA and the fiscal consolidations in Europe.

Government Purchases Multipliers based on Aggregate Data

Table 1 shows a sampling of estimates, grouped by method. Panel A shows estimates based on a variety of time series implementations and for a number of countries. Several are based on my updated estimates using best practices for computing multipliers that do not lead to biases. Panel B shows estimates based on New Keynesian DSGE models. The estimates show that for a variety of samples, identification methods, and countries, most of the estimates are around one or below. A few estimates are noticeably above one, such as the Ben Zeev and Pappa (2017) estimate, but they tend to be less precise and are not statistically different from one.³

On balance, the estimated multipliers are not very different across the two leading methods for identifying government spending shocks in time series nor for estimated New Keynesian DSGE models. With a couple of notable exceptions, most of the point estimates of multipliers are less than or equal to one.

³ These estimates are based on Ramey’s (2016) analysis using Ben Zeev and Pappa’s estimated news series.
Not shown in the table are numerous multiplier estimates based on key features of a country. For example, Iltetzki, Mendoza, Vegh (2010) estimate how multipliers change across various important features, such as fixed or flexible exchange rates. They find multipliers that vary between 0.1 on impact to 1.4 long-run (with a 90-percent confidence interval from around 0.75 to 2.1) for fixed exchange rates and from 0.1 to -0.7 for flexible exchange rates. Thus, the evidence they present suggests that the range of multipliers is much wider when one begins to distinguish by key country characteristic.

The results shown in Table 1 are for total government spending or government consumption. Earlier work by Aschauer (1989), Pereira and Flores de Frutos (1999), and others found high returns to public investment. There is surprisingly little recent aggregate evidence on multipliers for public investment. One example is from Iltetzki, Mendoza, Vegh (2010), who found multipliers for public investment that ranged between 0.4 in the short-run to 1.6 in the long-run in their panel of countries.

The bulk of the evidence suggests that government spending multipliers are probably one or below on average. However, the key question that has emerged is whether multipliers are higher during bad economic times. The key states studied by recent papers are recessions, periods of excess slack (typically measured by unemployment rates), constraints on the monetary policy accommodation, such as at the zero lower bound, and the ratio of public debt to GDP.

Consider first multipliers during recessions or periods of slack. The pioneering study on this question was by Auerbach and Gorodnichenko (2012), who used a nonlinear time series model in which the parameters changed across expansions and recessions. They reported a multiplier of 2.2 in recessions and -0.3 in expansions based on some simplifying assumptions about the state of the economy not changing after the shock. Various other studies have found
high multipliers during recessions (e.g. Auerbach and Gorodnichenko (2013), Fazzari, Morley, and Panovski (2015), Caggiano et al. (2015)). However, subsequent research has found many of the state dependent results to be very fragile to simple changes in specification or to improvements in the methods for computing the multipliers from the basic estimates (e.g. Alloza (2017), Owyang et al. (2013), Ramey and Zubairy (2018 and associated online appendix)). The more robust methods generally fail to produce multipliers above one during recessions or times of slack.

Perhaps these empirical results should not be so surprising, given the results of theory and quantitative models. Standard new Keynesian models do not predict higher multipliers during recessions. For example, Sims and Wolff (2018a) employ a medium-scale New Keynesian DSGE model with high order terms in the approximations and find that this otherwise standard model implies mildly procyclical multipliers. The only theoretical models that predict countercyclical markups are ones that include significant frictions. For example, Michaillat (2014) presents a stylized model with labor market frictions and finds that the aggregate employment effect of government hiring is countercyclical. However, the multipliers are always below one. Canzoneri et al. (2016) present a model with financial frictions that does generate sizeable, though fleeting, multipliers during recessions. They find significantly higher impact multipliers during recessions, near two, cut the cumulative multipliers fall below one after only a few quarters.

The situation is different with respect to periods when interest rates are near the zero lower bound or when monetary policy accommodates government spending increases (such as during WWII in the U.S.) Numerous New Keynesian DSGE models show that multipliers can be higher than one when monetary policy is constrained by the zero lower bound on interest
rates. At the zero lower bound, an increase in government spending provides extra stimulus through by increasing expected inflation, which lowers the real interest rate (Farhi and Werning (2016)). Calibrated models such as the ones analyzed by Christiano et al. (2011) and Coenen et al. (2012) can produce multipliers that range between 2 and 3 when the period of monetary accommodation is sufficiently long. Some recent empirical work has found some evidence of higher multipliers, ranging from 1.5 to 2.5 at the zero lower bound for Japan (Miyamoto et al. (2018)) and around 1.5 for historical samples in the U.S. (Ramey and Zubairy (2018)).

Finally, there is evidence that government spending multipliers may be negatively related to the public debt to GDP ratio. For example, Ilızetki et al. (2013) find that countries with a government debt to GDP ratio above 60 percent have an impact multiplier of 0 and a long-run multiplier of -3 (estimated less precisely but still statistically below 0).

In summary, most estimates of government spending multipliers for general categories of government spending for averages over samples are in the range of 0.6 to 0.8, or perhaps up to 1. The evidence for multipliers above one during recessions or times of slack is typically not robust. Some initial explorations suggest, though, that multipliers could be higher during zero lower bound periods.

**Tax and Transfer Multipliers based on Aggregate Data**

I now turn to the leading estimates of tax and transfer multipliers at the aggregate level. Table 2 shows the time series estimates and estimated DSGE models estimates for tax rate changes. Because many estimates of multipliers start out low on impact but then build, I report the cumulative multipliers for the horizon where they peak. I should also note that some of the
multipliers are calculated without allowing feedback from induced output changes to revenue whereas others allow for the dynamic feedback. The latter are noted in the table.

Most of the narrative method estimates are quite high, generally between -2 and -3. These are much higher (in absolute value) than the tax multipliers reported by Blanchard and Perotti (2002). As discussed above, their estimates were based both on their assumed elasticity of tax revenue to output and on their unusual way of computing multipliers. Barro and Redlick (2011) estimate multipliers around -1.1. It may be that their use of various approximations and constraints on dynamics account for their smaller estimate. On the other hand, Mountford and Uhlig’s (2009) estimates using sign restrictions are -5.

In contrast, the quantitative model estimates are much lower. Panel B of Table 2 shows that most quantitative model estimates yield multipliers that are below 1 in absolute value. Thus, there is a conflict between the time series estimates and the New Keynesian DSGE estimates.

There is a small literature on whether tax multipliers differ by the state of the economy. So far, this literature offers fairly uniform answers. Eskandari (2015) and Demirel (2016) find, using the Romer and Romer (2010) narrative tax shocks, that tax multipliers are greater during times of low unemployment than times of high unemployment. Alesina, Azzalini, Favero, Giavazzi, and Miano (2018) also find higher multipliers in expansions using their narrative of fiscal plans across OECD countries. These results are consistent with the one New Keynesian DSGE analysis of this issue. Sims and Wolff (2018b) obtain estimates of tax multipliers that are procyclical. For example, the capital tax multiplier is 1 in recessions and almost 2 in expansions.

There is very little work on the aggregate effects of transfers. Recently, Romer and Romer (2016) used changes in Social Security benefit increases to study the effects on macroeconomic variables. They found that permanent increases in benefits led to a roughly
equal rise in consumption in the short-run, but the effect dissipated quickly. Temporary increases in benefits had no significant effect on aggregate consumption. Coenen et al. (2012) studied general transfers and directed transfers across the various New Keynesian DSGE models used at policy institutions. They found that general transfers had multipliers between 0.2 and 0.6, with the higher ones occurring with monetary accommodation. In contrast, targeted transfers (to households that were financially constrained) yielded multipliers as high as 2 in some models when there was monetary accommodation.

In sum, most time series estimates of tax rate change multipliers indicate that they are very large, at least -2 to -3. This contrasts with the results from estimated New Keynesian DSGE models, where the multipliers are typically below 1 and never higher than 1.5. There is not much aggregate time series evidence for sizeable multipliers for temporary transfers, though calibrated New Keynesian models suggest they can be high if they are targeted and if monetary policy is accommodative.

**Multiplier Estimates based on Subnational Data**

As discussed earlier in this paper, one of the important innovations in the fiscal literature has been the application of applied microeconomics-type identification methods to the estimation of parameters of use for macroeconomics. These include studies of panels or cross-sections of U.S. states or provinces in other countries, as well as household-level estimates of marginal propensities to spend out of temporary transfers.

Chodorow-Reich (forthcoming) summarizes the panel and cross-section multipliers from individual studies, so I refer the reader to his tables. Many of the subnational multipliers for
government purchases, temporary tax rebates, and transfers lie between 1.5 to 2. Thus, they tend to be higher than the aggregate level estimates of multipliers.

As noted earlier, subnational multipliers are not the same as aggregate multipliers. The relationship between subnational multipliers and aggregate multipliers depends on many features, including how the spending is financed, whether there are spillovers across regions, whether there is a currency union, and whether the aggregate economy is at the zero lower bound. Nakamura and Steinsson (2014), Farhi and Werning (2016), and Chodorow-Reich (forthcoming) discuss a number of the theoretical considerations issues involved in drawing implications from subnational multiplier estimates to aggregate estimates. In some instances, the subnational multipliers are expected to be higher than the aggregate multipliers, whereas in other instances the subnational multipliers are expected to be lower than the aggregate multipliers. Thus, there is no general rule. A recent paper by Dupor and Guerrero (2017) conducts an interesting empirical investigation in which they directly compare estimates based on a state-level panel to those obtained when the state data are aggregated to the national level. They obtain similar multiplier estimates across the two data sets, though quite low, between 0 and 0.5.

**Multipliers in the Wake of the Financial Crisis**

A number of researchers and commentators have argued that the effects of the stimulus program in the U.S., the American Recovery and Reinvestment Act (ARRA), and the subsequent fiscal consolidations in European countries were much larger than indicated by multipliers during average times. They argue that the high unemployment rates and lower bound on interest rates combined to raise the multipliers.
As shown in the last sections, there is no robust evidence of higher multipliers during recessions or times of slack, for either spending or taxes. In fact, all studies of state dependence for tax multipliers find higher multipliers during expansions. However, there is evidence from historical periods in the U.S. and from Japan, as well as from New Keynesian models, that multipliers can be higher than one during periods of monetary accommodation such as the zero lower bound on interest rates. Thus, it is possible that multipliers could have been higher after the financial crisis.

Consider first the fiscal consolidations in Europe. Blanchard and Leigh (2013, 2014) presented evidence that countries that implemented bigger fiscal consolidations grew more slowly than forecasted by the IMF and other organizations. They concluded that forecasters assumed values of multipliers that were too small. Based on speculation about the size of the assumed multipliers, they concluded that the true multipliers during the crisis must have been above one. Recent work by Górnicka et al. (2018), though, gathered data on the assumed values of multipliers and found that they were very low, around 0.25. They calculated that the “true” ex post multipliers never exceeded one.

Górnicka et al.’s conclusions are consistent with some other analyses of the size of multipliers in the European fiscal consolidations. For example, Alesina, Favero, and Giavazzi (forthcoming) use their narrative data set of fiscal consolidation plans across OECD countries to study whether fiscal multipliers were greater in the immediate post-financial crisis years. They find no evidence that multipliers were greater. Thus, at this point, the evidence does not suggest that multipliers were larger than normal for the fiscal consolidations in Europe.

The American Recovery and Reinvestment Act (ARRA) was the leading stimulus program in the U.S. This program was a mix of spending and transfers to states and individuals.
As Table 3 shows, none of the New Keynesian DSGE models finds multipliers above unity for this program with the exception of one experiment by Coenen et al. (2012) with two years of monetary accommodation. While interest rates were indeed at the zero lower bound during those years, Swanson and Williams (2014) present evidence that yields on 1-year and 2-year treasury bills were unconstrained from 2008 to 2010, “suggesting that monetary policy and fiscal policy were about as effective as usual during this period.” (abstract).

In contrast, the cross-state estimates of the effects of the ARRA are typically much higher. Chodorow-Reich (forthcoming) presents an extremely valuable standardization and synthesis of the leading estimates of the effects of the ARRA on job creation across U.S. states. This literature emphasizes employment effects because the employment data have less measurement error than gross state product. He considers several of the leading instruments from the literature as well as their combination. The first-stage F-statistics are very high and suggest strong instruments. The results across the separate instruments are very similar and he cannot reject the over-identifying restrictions. Thus, the estimates are based on strong applied microeconomic methods. His cross-state natural experiment estimates indicate multipliers from 1.7 to 2 for gross state product and $50K per job-year created.

Building on Farhi and Werning’s (2016) theoretical analysis, Chodorow-Reich (forthcoming) argues that these subnational multipliers are lower bounds on the national multipliers during a liquidity trap. Thus, he argues that the ARRA multiplier was at least 2.

There is reason to suspect that the state-level estimates of the effects of the ARRA presented by Chodorow-Reich are probably overestimates for the national-level multipliers. The reason is that they answer two different questions. His cross-state estimates answer Question 1: “How much extra employment was induced in the average state by each $1 of ARRA spending
by the federal government?” The question relevant for the aggregate effects is Question 2: “How much extra aggregate employment was generated by each $1 of government spending induced by ARRA spending by the federal government?” Chodorow-Reich uses per capita values of spending and employment, so his cross-state estimates give equal weight to North Dakota and California, which is fine for answering Question 1. If there is heterogeneity in the treatment effects, however, the estimates will not give estimates that are nationally representative. Thus, the data need to be weighted to obtain nationally representative results. A second issue is the measure of government spending. Chodorow-Reich’s measure of spending is federal ARRA spending, which is appropriate for answering Question 1. However, as Leduc and Wilson (2017) show, the ARRA spending stimulated state and local spending more than dollar for dollar. Thus, multipliers that use only the ARRA transfers to the states will overestimate the multiplier per dollar spent across all levels of government.

Table 4 shows the effects of adjusting the employment response estimates to make them more suitable for answering Question 2. Table 4 shows three rows. The first row, which is from Table 1, column 4 of his paper shows Chodorow-Reich’s preferred estimates, which use all three of the leading ARRA instruments. The estimates are for job-years created for each $100K of the ARRA spending. The estimate of 2 implies that each $100K of ARRA spending creates 2 job years.

The second row of Table 4 shows the results of my re-estimating Chodorow-Reich’s model (using his replication files) but weighting by initial (Dec. 2008) state population. The point estimate to fall to 1.15, though the standard error is higher at 0.72. Thus, the point estimate

---

4 Most of the literature using cross-sectional estimates has used per capita estimates and has not weighted the estimates. However, Dupor and Mehkari (2016) started weighting the estimates and discovered that weighted estimates of the ARRA are much lower than unweighted estimates.
falls by over 40 percent once state observations are weighted by their population weights in order to make them representative of national data.

The third row of Table 4 shows the estimates when spending across the levels of government are substituted for the ARRA spending. I use his combination of instruments and I weight by initial state population. The jobs multiplier estimate is now 0.89 with a standard error of 0.45. Chodorow-Reich’s method for converting jobs multipliers to output multipliers is nearly one-for-one, so the 0.89 estimate also implies an output multiplier around 0.9.

Thus, once the cross-state estimates are made nationally representative and include all spending, they look very much like the aggregate estimates and lie below unity. Two important caveats about these adjusted estimates are in order, though. First, reweighting by population gives very large influence to just a few of the 50 states. Second, the great instrument relevance in Chodorow-Reich’s analysis disappears once I substitute state and local spending for his ARRA spending. It appears that the instruments that are so good at explaining ARRA spending are not very good at explaining total government spending in the state. Thus, it appears that the natural experiments exploited by the ARRA literature are rich enough to answer questions about the effects of ARRA spending in a typical state, but not to answer questions about the national effects of government spending induced by the ARRA.

In sum, a number of commentators and researchers have argued that multipliers may have been higher than usual after the financial crisis. Most of the evidence at this point suggests that they were not higher than usual.

**Conclusions**
The fiscal literature has made tremendous progress in the ten years since the start of the global financial crisis. The range of estimates for average multipliers has been reduced considerably. On average, government purchases multipliers are likely to be between 0.6 and 1 whereas tax rate change multipliers are likely to be between -2 and -3. However, there is still ongoing debate about specific contexts, such as the size of multipliers during “bad” times and the effects of other characteristics, such as exchange rate regimes.

I believe the literature would benefit from progress in three main areas. First, the literature needs to catch up to the current policy discussions by focusing more on the short-run and long-run effects of infrastructure investment. A few tantalizing studies at the aggregate and subnational levels suggest that these multipliers can be very large in some contexts. Second, researchers need to be careful about their implementation decisions. As I have shown, seemingly small changes, such as how multipliers are actually calculated, can make a big difference. Finally, researchers should continue to innovate along the same lines they have pursued in the last ten years, exploiting new data sets, extending theoretical models, and improving estimation techniques. As part of this innovation, researchers need to think more about the link between micro estimates and aggregate effects.
References


32


Table 1
Estimates of Government Spending Multipliers using Aggregate Data, No State Dependence

<table>
<thead>
<tr>
<th>Method/Sample</th>
<th>Multipliers</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Updated implementation of Blanchard-Perotti (2002) identified SVAR</td>
<td>0.6 to 0.8</td>
<td>The tax response is positive for the 1939q1-2015q4 period, but is essentially 0 for the later periods.</td>
</tr>
<tr>
<td>1939q1 – 2015q4</td>
<td>0.6 to 0.7</td>
<td></td>
</tr>
<tr>
<td>1947q1 – 2015q4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Military news shocks, local projections</td>
<td>0.6 to 0.8</td>
<td>Tax response is positive for 1939q1-2015q4 period.</td>
</tr>
<tr>
<td>Ramey-Zubairy (2018) military news</td>
<td>0.7 to 0.8</td>
<td>S.E. from 0.04 to 0.06</td>
</tr>
<tr>
<td>1889q1–2015q4</td>
<td>0.5 to 0.7</td>
<td>S.E. from 0.05 to 1</td>
</tr>
<tr>
<td>1939q1 – 2015q4</td>
<td>0.6 to 0.7</td>
<td>S.E. from 0.15 to 0.2</td>
</tr>
<tr>
<td>1947q1 – 2015q4</td>
<td>1.1 to 2</td>
<td>S.E. from 0.6 to 1.</td>
</tr>
<tr>
<td>Ben Zeev-Pappa (2017) news, 1947q1-2007q4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hall (2010), Barro and Redlick (2011) – based on regressions using annual</td>
<td>0.6 to 0.7</td>
<td>The Barro-Redlick analysis nets out effects of changes in tax rates.</td>
</tr>
<tr>
<td>defense spending.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mountford and Uhlig (2009), SVAR with sign restrictions</td>
<td>0.65</td>
<td>Deficit-financed increase in government spending.</td>
</tr>
<tr>
<td>Iltzetzki, Mendoza, Vegh (2013), BP identification in SVAR, quarterly data,</td>
<td>0.3 to 0.7</td>
<td></td>
</tr>
<tr>
<td>1960-2007,44 countries</td>
<td></td>
<td></td>
</tr>
<tr>
<td>High income countries</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corsetti, Meier, and Mueller (2012)</td>
<td>0.7</td>
<td>Based on unconditional model results reported in their Figure 1.</td>
</tr>
<tr>
<td>Leigh et al. (2010), Guajardo, Leigh and Pescatori (2014), 17 OECD countries,</td>
<td>0.3</td>
<td></td>
</tr>
<tr>
<td>1978 – 2009, narrative method identification of spending-based fiscal</td>
<td></td>
<td></td>
</tr>
<tr>
<td>consolidations</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alesina, Favero, and Giavazzi (forthcoming). Narrative analysis of austerity</td>
<td>0.3</td>
<td></td>
</tr>
<tr>
<td>plans, 16 OECD economies from 1978 - 2014.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### B. Estimated New Keynesian DSGE models

<table>
<thead>
<tr>
<th>Model Description</th>
<th>Multiplier</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cogan et al. (2010), estimated Smets-Wouters DSGE model on U.S. data</td>
<td>0.6 to 0.7</td>
<td>Based on my visual inspection of Figures 2, 3, and 4.</td>
</tr>
<tr>
<td>Coenen et al. (2012), large-scale macro models used by central banks and IMF, U.S. and Europe</td>
<td>0.7 to 1</td>
<td>Based on the two year cumulative multipliers shown in the upper left graph in their Figure 6.</td>
</tr>
<tr>
<td>Zubairy (2014), estimated medium scale DSGE model estimated on U.S. data.</td>
<td>0.7 to 1.05</td>
<td>Deficit financed, model features deep habits.</td>
</tr>
<tr>
<td>Leeper, Traum, and Walker (2017), estimated DSGE model on U.S. data.</td>
<td>0.7 to 1.36</td>
<td>Active monetary policy, Table 7</td>
</tr>
<tr>
<td>Sims and Wolff (2018a)</td>
<td>1.07</td>
<td>The multiplier above 1 is due to estimated complementarity of government spending with private consumption.</td>
</tr>
</tbody>
</table>
Table 2
Estimates of Tax Change Multipliers using Aggregate Data, No State Dependence
† denotes multipliers computed using the cumulative actual response of tax revenues or deficits in the denominator.

<table>
<thead>
<tr>
<th>Method/Sample</th>
<th>Peak cumulative multipliers within first 5 years.</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. Time Series Methods</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mountford and Uhlig (2009), SVAR with sign restrictions, U.S. data</td>
<td>-5†</td>
<td></td>
</tr>
<tr>
<td>Romer and Romer (2010), narrative series of tax changes unrelated to current economy, U.S. data, 1950 to 2007, dynamic single equation model or VAR</td>
<td>-2.5 to -3</td>
<td>The output effects take time to build.</td>
</tr>
<tr>
<td>Barro and Redlick (2011), historical annual U.S. data, tax rate shocks.</td>
<td>-1.1</td>
<td></td>
</tr>
<tr>
<td>Mertens and Ravn (2013, 2014), refinement of Romer and Romer series used in a proxy SVAR</td>
<td>-2.5 to -3†</td>
<td>The peak output effects occurs in the first 18 months.</td>
</tr>
<tr>
<td>Cloyne (2013), narrative, U.K.</td>
<td>-2.5</td>
<td></td>
</tr>
<tr>
<td>Hayo-Uhl (2013), narrative, Germany</td>
<td>-2.4</td>
<td></td>
</tr>
<tr>
<td>Alesina, Azzalini, Favero, Giavazzi, Miano (2018) Narrative analysis of austerity plans, 16 OECD economies from 1978 – 2014, taxed based consolidations</td>
<td>Based on static primary surplus</td>
<td>-1 to -1.6</td>
</tr>
<tr>
<td>Based on actual response of primary surplus</td>
<td>-2.3 to -3.7†</td>
<td></td>
</tr>
</tbody>
</table>

† denotes multipliers computed using the cumulative actual response of tax revenues or deficits in the denominator.
### B. Estimated New Keynesian DSGE models

<table>
<thead>
<tr>
<th>Model</th>
<th>Fiscal Policy</th>
<th>Consumption tax</th>
<th>Labor tax</th>
<th>Corporate income tax</th>
<th>Labor tax</th>
<th>Capital tax</th>
<th>Capital tax</th>
<th>Steady-state multipliers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coenen et al. (2012), large-scale macro models used by central banks and IMF, U.S. and Europe. Two-year cuts in tax, no monetary accommodation</td>
<td></td>
<td>-0.2 to -0.4</td>
<td>-0.2 to -0.4</td>
<td>0 to -0.15</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zubairy (2014)</td>
<td></td>
<td>-0.7 to -1</td>
<td>-0.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sims-Wolff (2018b) – medium scale New Keynesian DSGE model that allows for higher order terms.</td>
<td></td>
<td>-0.6</td>
<td>-1</td>
<td>-1.5</td>
<td></td>
<td></td>
<td></td>
<td>Steady-state multipliers</td>
</tr>
</tbody>
</table>
### Table 3
Multipliers for the American Recovery and Reinvestment Act (ARRA)

<table>
<thead>
<tr>
<th>Method/Sample</th>
<th>Peak cumulative multipliers within first 5 years.</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cogan et al. (2010)</td>
<td>0.6 to 0.7</td>
<td></td>
</tr>
<tr>
<td>Coenen et al. (2012), large-scale macro models used by central banks and IMF, U.S. and Europe</td>
<td>0.3 to 0.5</td>
<td>From Figure 7. These are the peak instantaneous multipliers.</td>
</tr>
<tr>
<td></td>
<td>0.4 to 0.6</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.5 to 1.8</td>
<td></td>
</tr>
<tr>
<td>Drautzburg and Uhlig (2015), medium scale New Keynesian DSGE model, with ZLB, credit constraints</td>
<td>0.5</td>
<td>Multipliers become negative in the long-run because of the necessary increase in taxation.</td>
</tr>
<tr>
<td>Chodorow-Reich (forthcoming), based on cross-state estimates and theoretical arguments about the relationship between subnational and national multipliers at the ZLB.</td>
<td>1.7 to 2</td>
<td></td>
</tr>
<tr>
<td>GSP multiplier</td>
<td>2 job-years per $100K</td>
<td></td>
</tr>
<tr>
<td>Cost per job year</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 4. Conversion of Chodorow-Reich Estimates to National Estimates

<table>
<thead>
<tr>
<th></th>
<th>Cumulative Employment Multiplier Estimates – Number of Job Years Created per $100K of ARRA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chodorow-Reich headline estimates (his Table 1, column 4)</td>
<td>2.01 (0.59)</td>
</tr>
<tr>
<td>Weighted estimates (using Dec. 2008 population of state)</td>
<td>1.15 (0.72)</td>
</tr>
<tr>
<td>Weighted estimates</td>
<td></td>
</tr>
<tr>
<td>Estimates based on total spending, including induced spending by states</td>
<td>0.89 (0.45)</td>
</tr>
</tbody>
</table>

Notes. Estimates presented in the last two rows are the author’s estimates, based on Chodorow-Reich’s programs and data.
Figure 1
Estimated Impulse Response Functions for a Shock to Government Purchases

Source: Author.
Note: Estimated impulse responses based on SVAR estimates using the Ramey-Zubairy (2018) data. The sample is quarterly, 1939:1 – 2015:4. The SVAR contains five endogenous variables: log real total government spending per capita, log real GDP per capita, log real federal tax receipts per capita, the 3-month Treasury bill interest rate, inflation measured as the log change in the GDP deflator. The SVAR includes four quarterly lags of variables. The shock to government spending is identified using Blanchard and Perotti’s (2002) method, which orders government spending first. The shaded area shows the 95-percent confidence bands.
Figure 2
Alternative Definitions of Multipliers

Source: Author.
Note: The dotted and solid lines show multipliers calculated based on the log impulse responses shown in Figure 1. Let $y(j)$ denote the value of the impulse response of log(GDP) at horizon $j$ and $g(j)$ denote the impulse response of log(government spending).

Quasi-multiplier ($h$) = \((\frac{\bar{Y}}{\bar{G}}) \cdot \frac{y(h)}{g(0)}\), where \((\frac{\bar{Y}}{\bar{G}})\) is the sample average of GDP to government spending, equal to 4.78.

Present value (PV) cumulative multiplier ($h$) = \((\frac{\bar{Y}}{\bar{G}}) \cdot \frac{\sum_{j=0}^{h} \beta^j y(j)}{\sum_{j=0}^{h} \beta^j g(j)}\) for log SVAR, where $\beta = \frac{1}{1+r}$, where $r$ is the real interest rate, assumed to be 3.6% on an annual basis.

The line with diamonds shows the multiplier using the Gordon-Krenn transformation, where

Present value (PV) cumulative multiplier ($h$) = \(\frac{\sum_{j=0}^{h} \beta^j Y_{GK}(j)}{\sum_{j=0}^{h} \beta^j G_{GK}(j)}\), where $Y_{GK}$ and $G_{GK}$ are the responses of the Gordon-Krenn transformed variables and the estimates are from the alternative SVAR that uses those variables.