Valerie Ramey Discussion of "Fiscal Policy" by Alan Auerbach

November 17, 2017

"Fiscal Policy" by Alan Auerbach is a must-read for anyone contemplating current fiscal challenges facing developed countries as well as the role of fiscal policy in stabilizing the economy. The article is filled with deep insights on numerous facets of the issues involved. Alan's article sheds light on the issues involved in four major areas: (i) Fiscal rules; (ii) Potential for stabilization policy; (iii) Fiscal policy in a low-interest rate environment; and (iv) Coordination between monetary and fiscal policy. Among the many important insights in the in the article is Alan's argument that basing fiscal rules on observable public debt while ignoring the more important unfunded liabilities involved in entitlement programs misses the main source of potential future debt issues. In my comments, I will focus on area (ii) – the potential for fiscal policy as a stabilization policy - as that is the area with which I am most familiar.

I will begin by offering my assessment of the U.S. evidence. I summarize my views in the following three statements:

1. Spending stimulus multipliers are probably less than one in most instances.
2. The evidence on infrastructure multipliers is mixed.
3. The strongest, most robust evidence is that tax cuts have the biggest multipliers.

To support these statements, I will begin by considering the evidence related to spending multipliers. Aggregate spending multipliers in U.S. are typically estimated to be just below unity, both in post-WWII data and in historical data covering the 20th Century (e.g. Ramey and Zubairy (forthcoming)). In an important paper, Auerbach and Gorodnichenko (2012) investigated whether those averages could be masking important variations in multipliers over time, and in particular,

during recessions versus expansions. They found much higher multipliers during recessions. Let us look at those estimates more closely.

Auerbach and Gorodnichenko (2012) (AG-12) report baseline estimates of 5-year multipliers that are 2.2 in recessions and -0.3 in expansions. AG-12 make two key assumptions when they construct these estimates. First, they assume that the recession lasts at least five years, or 20 quarters. Second, they assume that government spending cannot get the economy out of a recession. When AG-12 relax the second assumption, they find multipliers in recessions in the range from 1 to 1.5. However, even these multipliers are affected by their first assumption, which is that recessions last at least five years. In their sample, the median length of a recession is only three quarters and, by their definition, even the Great Recession lasted only two years. Making an assumption that is so contrary to the experience in their estimation sample turns out to have big effects on their multiplier estimates. In particular, because recessions are usually short, the estimates imply rising output growth in the future as the economy recovers from the slow growth of the recession. When one counterfactually assumes that the economy remains in recession for years, one cumulates these faster-than-current growth forecasts so that they predict that output will continuing climbing indefinitely relative to trend.

An alternative method, which Auerbach and Gorodnichenko (2013) introduced in their study of multipliers in the OECD, is Jordà's (2005) local projection method, which produces estimates of the effects of government spending based on the average length of recessions and expansions over the sample. This method does not allow one to make assumptions that are counterfactual to the data analyzed. Auerbach and Gorodnichenko never applied this method to the U.S. data, though. In Owyang, Ramey and Zubairy (2013) and in Ramey and Zubairy (forthcoming), my co-authors and I use the Jordà local projection method on U.S. data, both

historical and AG-12's post-WWII sample. In both cases, we find multipliers that are below one even in recessions or when the unemployment rate is high, around 0.8 for both samples. These estimates are obtaining using a military news series as shock; when we instead use Blanchard and Perotti's (2002) identification method, we find even lower multipliers in recessions (thought they are greater than the near-zero multipliers estimated during expansions). Thus, applying Auerbach and Gorodnichenko's currently-favored Jordà technique to the U.S. data produces multipliers estimates that are below one even in recessions. In sum, the original findings of high multipliers during recessions in the U.S. are fragile.

In our forthcoming paper, Zubairy and I also investigate the size of multipliers during zero lower bound periods. When we use the entire historical sample, we find no evidence of elevated multipliers during zero lower bound periods. However, if we exclude the period of WWII rationing we find some evidence of multipliers that are around 1.5 during zero lower bound periods. These results hold when we use military news identification, but not the Blanchard and Perotti identification. In follow-up work on the Japanese economy, Miyamoto, Nguyen, and Sergeyev (forthcoming) also find evidence of multipliers around 1.5, or even higher, during Japan's zero lower bound decades. Thus, there is some evidence of higher multipliers during zero lower bound periods for some samples.

To summarize, aggregate estimates suggesting multipliers above one during recessions or slack times tend to be fragile. The most robust estimates are those that are just below one. On the other hand, there is mounting evidence at this point that zero lower bound periods might be associated with multipliers that are greater than one.

Numerous papers have estimated multipliers using cross-state or cross-region data or constructed multipliers using marginal propensities to consume estimated using individual level

data. Identifying exogenous shocks at the household or state level is often easier because of handy natural experiments. Interestingly, the individual-level estimates often imply high marginal propensities to consume (MPCs) and state-level estimates often imply high multipliers (1.5 to 2 for GDP multipliers or around $50K per job-year for employment multipliers). However, exporting these cross-sectional estimates to the aggregate level is not straightforward, as numerous papers have argued.

To determine whether it makes sense to use cross-sectional estimates directly as aggregate estimates, it is useful to conduct simple plausibility tests. I will present two here. The first is directly from Sahm, Shapiro, Slemrod (2012) and regards MPCs and the second is my own based on the cross-state estimates of the effects of the stimulus during the Great Recession.

Consider first Sahm, Shapiro, and Slemrod's (2012) plausibility test of some recent individual-level estimates of the marginal propensity to consume out of temporary tax rebates. In 2008, a temporary tax rebate ranging from $300 to over $1800, depending on household size was enacted in February 2008 and disbursed mostly from April through July. Fortunately for economic researchers, the disbursement timing was randomized by Social Security numbers. Parker, Souleles, Johnson and McClelland (2013) worked with the BLS to add a supplement to Consumer Expenditure Survey in order to study the effects of the rebates on consumer expenditures. They looked at a variety of categories but one striking finding, and the one focused on by Sahm et al. (2012), was that on average consumers spent 40 percent of their rebate check on motor vehicle purchases. Sahm et al. (2012) use that estimate to conduct a simple counterfactual: ignoring any general equilibrium multipliers or price changes, what would motor vehicle expenditures have

been in 2008 had there been no rebate? They show the results in Table 14 at the end of their paper. I graph the results here because it is easier to see the nature of the counterfactual path.[1]

Figure 1. Actual and Counterfactual Expenditures on Motor Vehicles
(Counterfactual estimates from Sahm, Shapiro, and Slemrod (2012), Table 14)

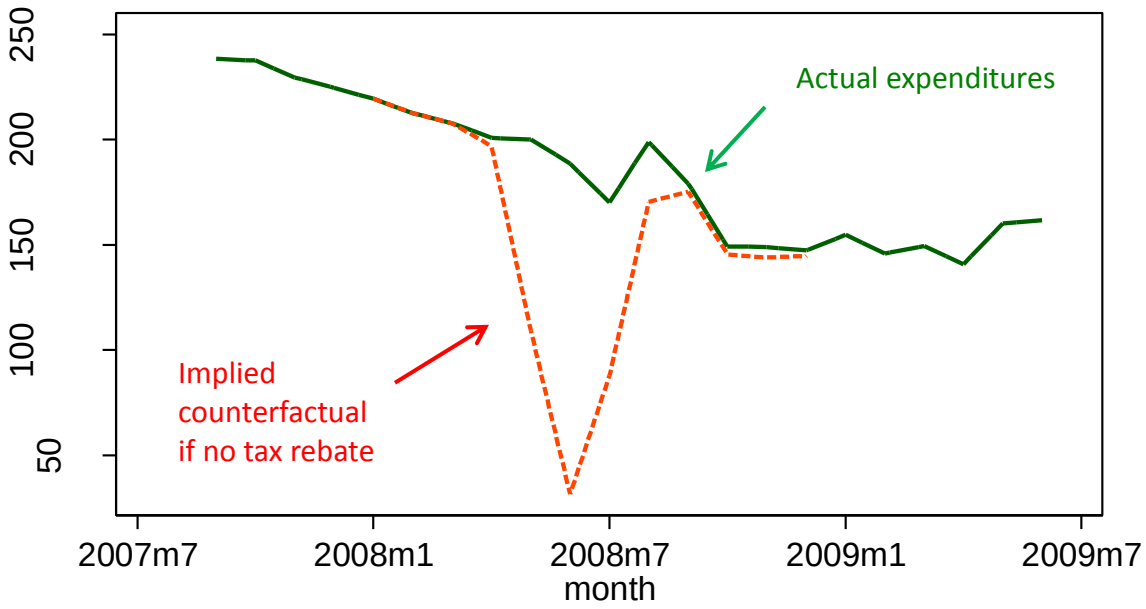


Figure 1 shows the striking counterfactual. If one applies the marginal propensity to spend the rebate on motor vehicles estimated by Parker et al. (2013) to the aggregate data, it implies that *most* of the spending on motor vehicles in Summer 2008 was due to spending induced by the tax rebate. It implies that in the absence of the tax rebate, motor vehicle expenditures would have collapsed from a $208 billion annual rate in March 2008 to only $31 billion annual rate in June 2008. Moreover, it implies that in the absence of a rebate, motor vehicle spending would have rebounded sharply when Lehman Brothers collapsed. This counterfactual is preposterous and

[1] The only difference between the numbers in the graph and those in their table is that I use current vintage motor vehicle expenditure data that extends before and after 2008 in order to give context

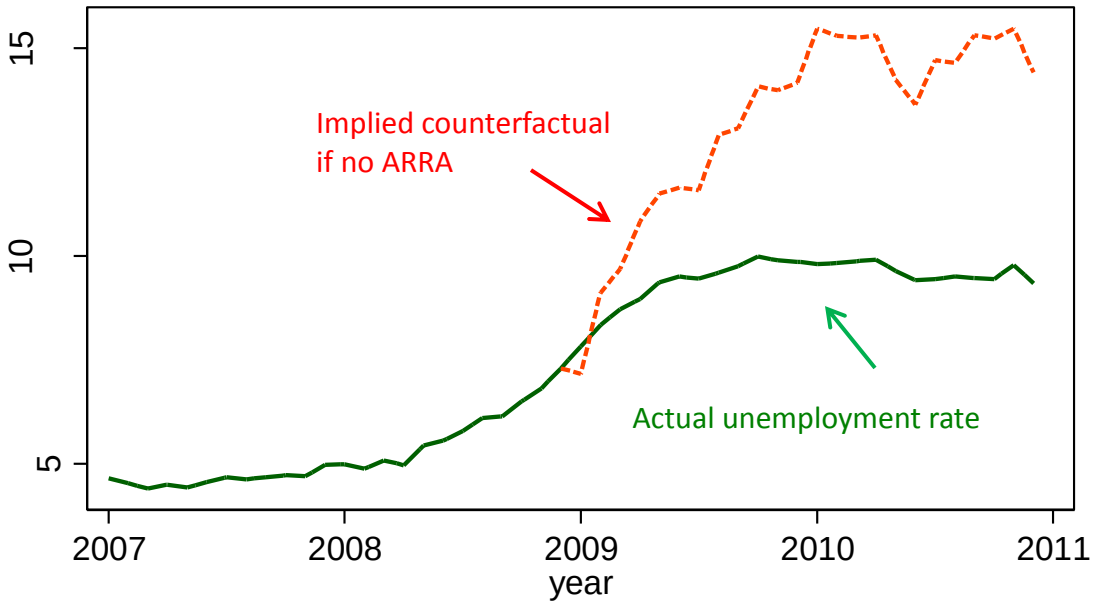
serves as a cautionary tale about the complicated link between cross-sectional estimates and aggregate estimates.

The second counterfactual I conduct is related to the effects of the American Recovery and Reinvestment Act (ARRA) stimulus spending. A recent paper by Chodorow-Reich (2017) synthesizes and standardizes the various estimates in the literature on the effect of the ARRA across states. His preferred estimate is that each $50K created one job that lasted a year (a "job-year"). Stated as a GDP multiplier, he estimates multipliers around 2. Importantly, Chodorow-Reich (2017) goes on to argue, based both on a New Keynesian model with a zero lower bound and some back of the envelope calculations, that the cross-state estimates of multipliers are *lower bounds* on the aggregate multiplier during zero lower bound times.

In the spirit of Sahm et al.'s counterfactual exercise, I conduct a counterfactual about the ARRA using Chodorow-Reich's estimates. In particular, I use his Figure B.1 estimates of the impulse response of employment to the passage of the stimulus bill for the period December 2008 through December 2010. I also use his estimate that $600 billion of the ARRA had been spent by December 2010. I then calculate what his estimates imply for the lower bound of induced monthly employment at the aggregate level if we believe that the cross-state estimates are lower bounds on the aggregate. I transform this into a counterfactual unemployment rate by adding the induced employment to the actual number unemployed. The result is shown in Figure 2.

Figure 2 shows that applying Chodorow-Reich's cross-state ARRA employment estimates to the aggregate implies that the unemployment rate would have risen to 15.5 percent had the ARRA not been passed. The actual unemployment rate rose 2.7 percentage points from December 2008 to its peak of 10 percent. The counterfactual path implied by Chodorow-Reich's

Figure 2. Actual and Counterfactual Unemployment Rate
(Counterfactual rate based on the author's estimates)



estimates show an unemployment rate that would have risen not 2.7 percentage points, but 8 percentage points, to its peak of 15.5 percent. Thus, applying his cross-state estimates to the aggregate imply that the unemployment rate would have risen by three times more in those two years than it actually did. While not as outlandish as the Sahm et al. counterfactual, this counterfactual does strain plausibility.

To summarize my points for government spending multipliers for the U.S., most estimates are below one and the few estimates that are above one are typically not robust. Furthermore, estimates of multipliers at the cross-sectional level produce implausible results when directly applied to the aggregate level. This is not to say that an increase in government spending does not raise GDP, but only that it raises GDP by less than the rise in government spending.

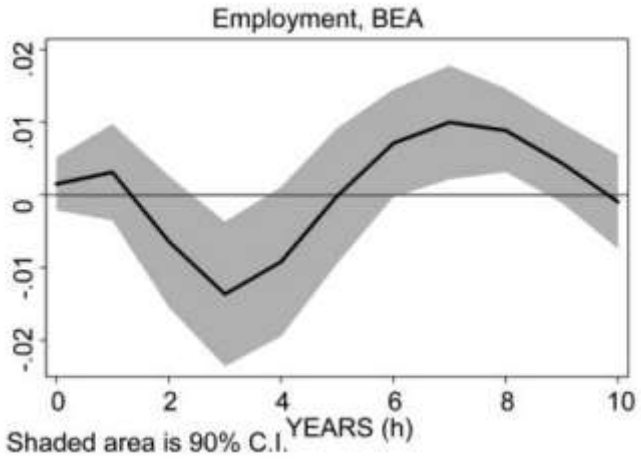
My second point is about infrastructure spending. In contrast to government spending that is not used for public capital, infrastructure spending has the potential to stimulate the economy not only through standard Keynesian multipliers, but also through a supply-side effect. Recent

commentators have advocated increased infrastructure spending for this reason. Little is certain, however, about the size of the infrastructure multiplier. Studies of the effects of aggregate spending projects, such as the U.S. interstate highway system, are faced with the challenge of identifying which part of aggregate output is affected by the one big wave of highway spending. Fernald's (1999) study of the differential effects of the interstate highway program on particular transportation-dependent industries suggest large effects. However, these effects are *relative* industry effects, not aggregate effects, and as Fernald points out, a large effect of the initial highway program is no guarantee of such a large effect of subsequent projects where diminishing returns might set in.

A very useful recent cross-sectional contribution is by Leduc and Wilson (2013), who analyze the effects of highway construction across states in more recent decades. They are very careful about every detail of the nature of and timing of the announcements and spending. There are aspects of their results that raise questions about the ability of infrastructure spending to stimulate the economy in the short and medium run, however. Consider for example their estimates of the effects of infrastructure spending on employment in a state. I focus on employment because the results are most stark for that variable. Figure 3 is copied from the upper right graph of Figure 4 of their paper. Note that the effects are essentially 0 on impact in year 0, then become negative through year 5 and finally become positive in years 6 through 9 before returning to normal at year 10. Visually integrating over the graph it appears that the cumulative effect on employment through year 10 might even be negative.

If most estimates of government spending multipliers are low, then is there no role for fiscal policy for stimulating the economy? Not necessarily. The largest, most robust multipliers

Figure 3. The Effects of State Highway Spending on State Employment
(Copied from one graph of Figure 4 of Leduc and Wilson (2013))



appear to be for tax cuts, particularly those involving cuts in tax rates. Romer and Romer (2010) use narrative methods to identify U.S. tax changes for reasons unrelated to the current state of the economy and find multipliers as high (in magnitude) as -3 (the minus sign is because a decrease in taxes raises GDP). Cloyne (2013) finds similar results for the U.K. Mertens and Ravn (2014) show the robustness of Romer and Romer's result for the U.S., estimating multipliers between -2.5 and -3 using other methods, and explain the lower multiplier estimate obtained by Blanchard and Perotti (2002) by a restriction they impose in order to identify tax shocks. In Mertens and Ravn (2013), they split the Romer and Romer series into changes in personal income tax rates and corporate tax rates. They estimate that a one percentage point cut in the average personal income tax rate leads real GDP to increase 1.4 percent on impact and up to 1.8 percent after three quarters. This implies a personal income tax multiplier of up to -2.5. A one percentage point cut in the corporate income tax rate leads real GDP to rise by 0.4 percent on impact and 0.6 percent after a year. Notably, Mertens and Ravn cannot calculate the tax multiplier for a corporate income tax cut because they estimate that corporate income tax cuts *do not lower tax revenues*. That is, the

stimulus effect of a corporate income tax cut on measured GDP is so large that tax revenues do not fall.

In sum, the strongest, most robust results in the literature are those for tax changes. The estimates imply a large potential role for tax cuts as a stimulus to the economy. Thus, tax policy should receive at least equal billing with government spending policy when discussing ways that fiscal policy can stimulate the economy.

References cited:

Auerbach, Alan and Yuriy Gorodnichenko. 2012. "Measuring the Output Responses to Fiscal Policy." *American Economic Journal: Economic Policy* 4 (2): 1-27.

Auerbach, Alan and Yuriy Gorodnichenko. 2013. "Fiscal Multipliers in Recession and Expansion" *Fiscal Policy After the Financial Crisis*, eds. Alberto Alesina and Francesco Giavazzi, University of Chicago Press.

Blanchard, Olivier, and Roberto Perotti, "An Empirical Characterization of the Dynamic Effects of Changes in Government Spending and Taxes on Output," *Quarterly Journal of Economics*, 117 (2002), 1329-1368.

Chodorow-Reich, Gabriel. 2017. "Geographic Cross-Sectional Fiscal Spending Multipliers: What Have We Learned?" Harvard working paper.

Cloyne, James, "Discretionary Tax Changes and the Macroeconomy: New Narrative Evidence from the United Kingdom," *The American Economic Review*, Volume 103, Number 4, June 2013, pp. 1507-1528(22).

Jordà, Òscar. 2005. "Estimation and Inference of Impulse Responses by Local Projections." *American Economic Review* 95 (1): 161-82.

Leduc, Sylvain and Daniel Wilson. "Roads to Prosperity or Bridges to Nowhere? Theory and Evidence on the Impact of Public Infrastructure Investment," (*NBER Macroeconomics Annual* 27, no. 1 (2012), pp. 89-142.

Mertens, Karel, and Morten O. Ravn. 2013. "The Dynamic Effects of Personal and Corporate Income Tax Changes in the United States." *American Economic Review*, 103(4): 1212-47.

Mertens, Karel, and Morten O. Ravn., "A Reconciliation of SVAR and Narrative Estimates of Tax Multipliers," *Journal of Monetary Economics* 68 (2014): S1-S19.

Miyamoto, Wataru and Nguyen, Thuy Lan, "News Shocks and Business Cycles: Evidence from Forecast Data," forthcoming *American Economic Journal: Macroeconomics.*

Owyang, Michael T., Valerie A. Ramey, and Sarah Zubairy. "Are Government Spending Multipliers Greater during Periods of Slack? Evidence from Twentieth-Century Historical Data." *The American Economic Review* 103.3 (2013): 129-134.

Parker, Jonathan A, Nicolas S. Souleles, David S. Johnson, Robert McClelland, "Consumer Spending and the Economic Stimulus Payments of 2008," *The American Economic Review,* 103 (October 2013): 2530-2553.

Ramey, Valerie A. and Sarah Zubairy, "Government Spending Multipliers in Good Times and in Bad: Evidence from 20th Century Historical Data," forthcoming *Journal of Political Economy*.

Romer, Christina D., and David H. Romer, "The Macroeconomic Effects of Tax Changes: Estimates Based on a New Measure of Fiscal Shocks," *American Economic Review,* 100 (June 2010): 763-801.

Sahm, Claudia R., Matthew D. Shapiro, Joel Slemrod, "Check in the Mail or More in the Paycheck: Does the Effectiveness of Fiscal Stimulus Depend on How it is Delivered?" *American Economic Journal: Economic Policy,* 4 (August 2012):216-250.