Follow-up on July 12, 2014 EFG Discussion about Government Spending Multipliers

By Valerie A. Ramey and Sarah Zubairy

July 29, 2014

This note is a follow-up to the July 12, 2014 NBER EFG discussion between Valerie Ramey and Sarah Zubairy, who presented their paper "Government Spending Multipliers in Good Times and in Bad: Evidence from U.S. Historical Data," and Yuriy Gorodnichenko, who was the discussant.

Two key issues were raised in Yuriy's discussion. First, he showed results from an alternative specification that delivered higher multipliers in recessions, and second, he questioned the first stage F-statistics of the military news variable relative to the Blanchard-Perotti identified shock. We were not able to resolve the issues on the spot because they depended on details of specification that were not clear to us from the slides that Yuriy had sent us. He has since sent us the programs so we were able to inspect the details. We now discuss each of these issues. Based on our analysis, we have come to the following conclusions:

1. Yuriy's high multipliers were due both to the unconventional way he translates impulse responses into multipliers and to details of his implementation of the Blanchard and Perotti identification. Thus, his results **do not** indicate a lack of robustness to the other details of our specification, such as the number of lags and whether variables are in first-differences.

2. Yuriy's method for calculating F-statistics for each horizon provides useful statistics about the information value of the shocks. Using his method, however, we show that the military news shock is more informative at some horizons whereas the Blanchard-Perotti identified shock is more informative at other horizons.

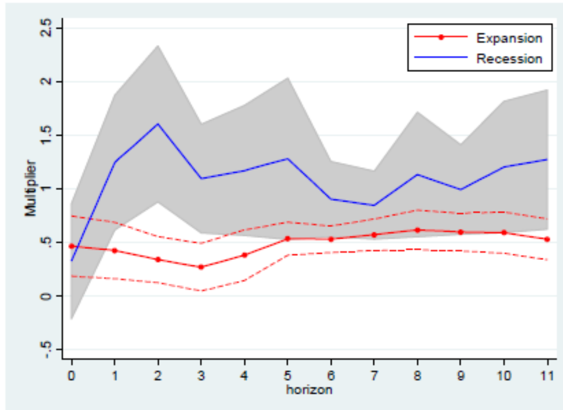
We now discuss the details of each.

(1) How did Yuriy Gorodnichenko obtain higher multipliers during recessions?

Figure 1 shows a copy of Slide 41 of Yuriy's discussion (the entire set of slides is available here). This slide suggests that multipliers during recessions are as high as 1.5 at some horizons during the first twelve quarters. Table 3 of our paper showed a variety of robustness checks, including one using the Blanchard-Perotti identification. We had reported two-year multipliers (equivalent to horizon 7 on the graph) of 0.73 when unemployment was high and 0.51 when unemployment was low.

Figure 1. Slide 41 of Yuriy's Discussion

MULTIPLIERS: BLANCHARD-PEROTTI



Spec: IV implementation, include more lags, normalize by potential GDP, controls include variables in *growth rates* rather than levels.

These estimates are similar to the Auerbach-Gorodnichenko results.

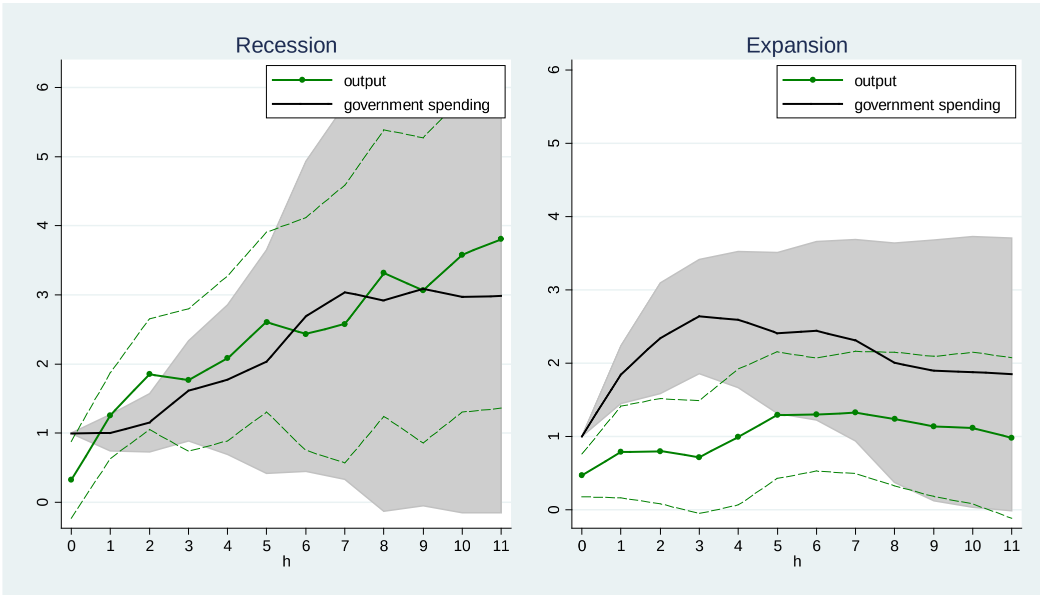
Our inspection of Yuriy's program revealed two key details in his specification that resulted in higher multipliers. These two details were: (A) the method he used to compute the multipliers; and (B) his implementation of the Blanchard-Perotti identification. The details that he claimed mattered – adding more lags, normalizing by potential GDP, and including variables in growth rates – had little or no effect. We now discuss the issues with his calculation of multipliers and his implementation of Blanchard-Perotti.

A. Yuriy's method for computing "multipliers"

Blanchard and Perotti (2002) calculated their multipliers by comparing the *peak* response of output to the *initial* government spending shock. Auerbach and Gorodnichenko (2012, 2013) also report similar multipliers. Mountford and Uhlig (2009), Uhlig (2010), and Fisher and Peters (2010) argued that the multipliers relevant for policymakers are instead ones that compare the *cumulative* output response to the *cumulative* path of government spending. We agreed and therefore reported the cumulative multipliers through various horizons in our paper. In contrast, Yuriy computes multipliers as the ratio of the impulse responses at particular *points in time.* While this may be a useful description of the impulse response functions, it is not a conventional multiplier because it does not properly account for the full cost of a government spending shock.

In his discussion, Yuriy implemented an IV procedure which produced estimates only of the ratio of the output response at horizon *h* to the government spending response at horizon *h*. To see how the cumulative multipliers behave for Yuriy's specification, we estimated his specification in two steps (as we do in our paper) so that we could extract the individual impulse responses. Figure 2 shows the impulse response functions in each state.

Figure 2. Impulse Responses from Yuriy's Specification



Denote the impulse response function of output at horizon *h* by $\beta_{Y,h}$ and the response of government spending at horizon *h* by $\beta_{G,h}$. Specifically, Yuriy's multipliers are the ratio of the output response to the government spending response *at each horizon,* given as follows:

$$\frac{\beta_{Y,h}}{\beta_{G,h}}$$

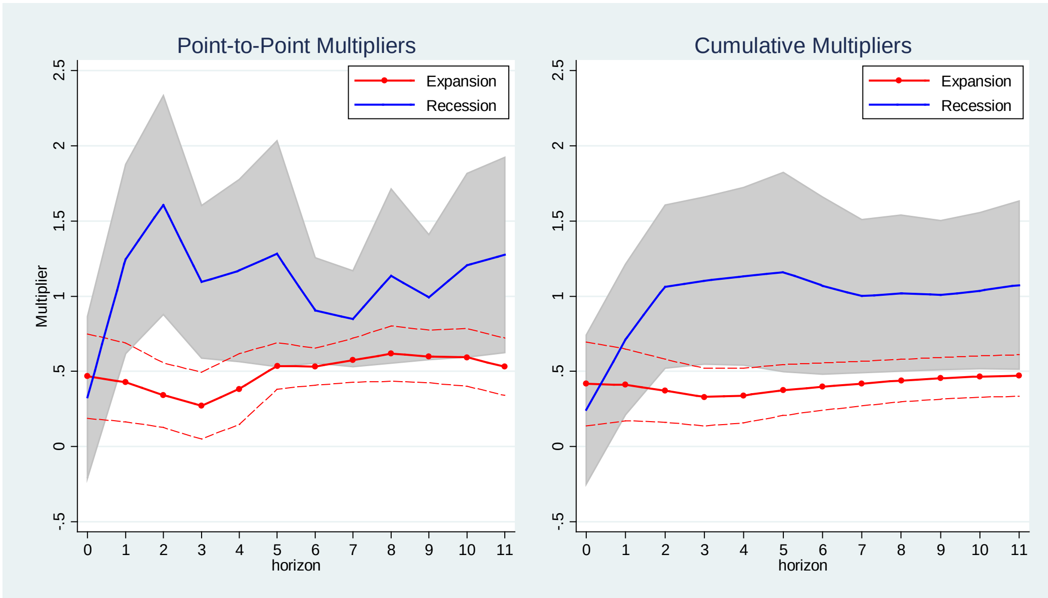
We instead calculate multipliers using the cumulative response method, that is, the integral under the impulse responses up to horizon *h*, as follows:

$$\frac{\sum_{i=0}^{h}\beta_{Y,i}}{\sum_{i=0}^{h}\beta_{G,i}}$$

Figure 3 below shows multipliers calculated the two ways. The point-to-point graph on the left contains the multipliers that Yuriy showed. He noted that one of the responses was just over 1.5, which was closer to the range reported in Auerbach-Gorodnichenko (2012). The right hand graph shows the cumulative multipliers, calculated using the integral under the impulse responses.

The cumulative multiplier, which incorporates the fact that output rose less than government spending on impact, does not increase as much as the point-to-point multiplier in the recession state. It hits a value slightly above 1 at horizon 2 and stays around there through horizon 11. The differences between the cumulative multipliers in recessions and expansions starting at horizon 2 are statistically significant at standard levels of significance. However, the difference between multipliers is not because the multiplier is so high in recessions but rather because it is so low in expansions. The recession multipliers are close to one, and the standard error bands encompass the common range of many non-state dependent estimates.

Figure 3. Yuriy's Specification: Comparison of Multiplier Calculations



(B) Details of Yuriy's Implementation of Blanchard-Perotti

The Blanchard and Perotti (2002) (BP) method of identifying government spending shocks uses a Choleski decomposition in a system in which government spending is ordered first. Blanchard and Perotti (2002) use the logarithms of real per capita government spending, GDP, and net taxes in their VARs. They estimate two specifications – one in log levels with a polynomial in trends and one in first-differences. The first difference specification identifies the shock to government spending to be the part of the current growth rate of government spending that is orthogonal to four lags of the growth of government spending, GDP and net taxes. In our robustness checks, we implemented the BP identification as the shock to the log of real per capital government spending, orthogonalized with respect to lags of the logs of government spending and GDP.

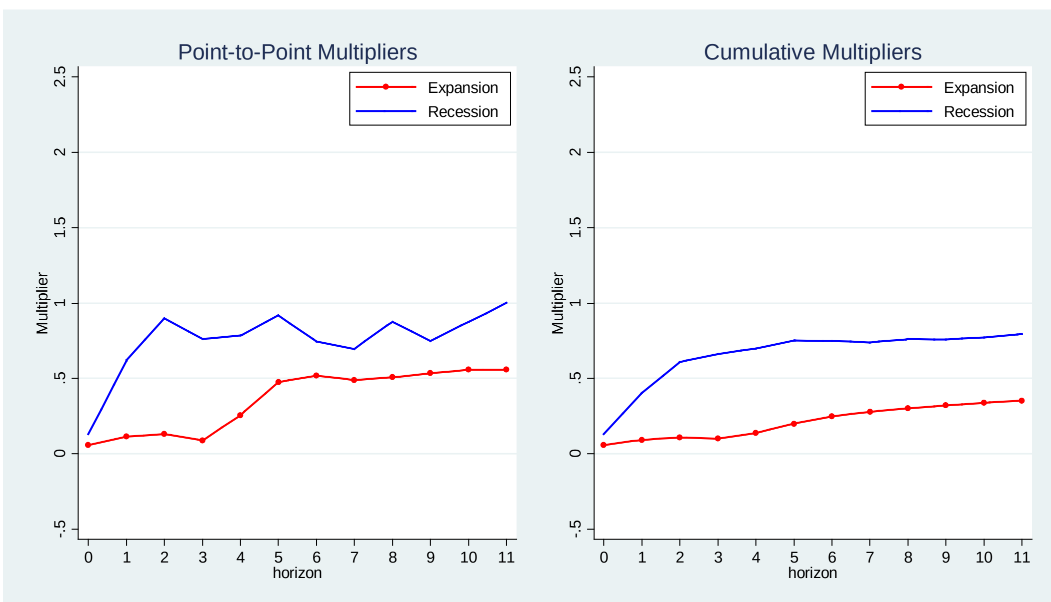
Yuriy's implementation orthogonalizes a different variable to create the shock. In particular, he uses the following variable to create the shock:

$$\frac{\Delta G_t}{Potential\ Y_t}$$

Here, *G* is the level of real government spending and *Y* is real GDP. Yuriy uses a Hodrick-Prescott filter to create potential GDP. The shock is identified as the part of this normalized government spending variable that is orthogonal to the four lags of itself, $\Delta\ ln(g_t)$ and $\Delta\ ln(y_t)$, and a quartic time trend, where the lower case letters indicate per capita real government spending and output . This shock variable is analogous to the Hall-Barro-Redlick dependent variables we advocate as a solution to the "units" problem. However, because units don't matter for the instrument, one does not need to perform this transformation. Yuriy also changes the formulation of both the dependent variables and the shock so that rather than lagged GDP in the denominator, he uses an estimate of potential GDP based on an HP filter. We have concerns about including an HP filter trend in an identification scheme that relies on timing.

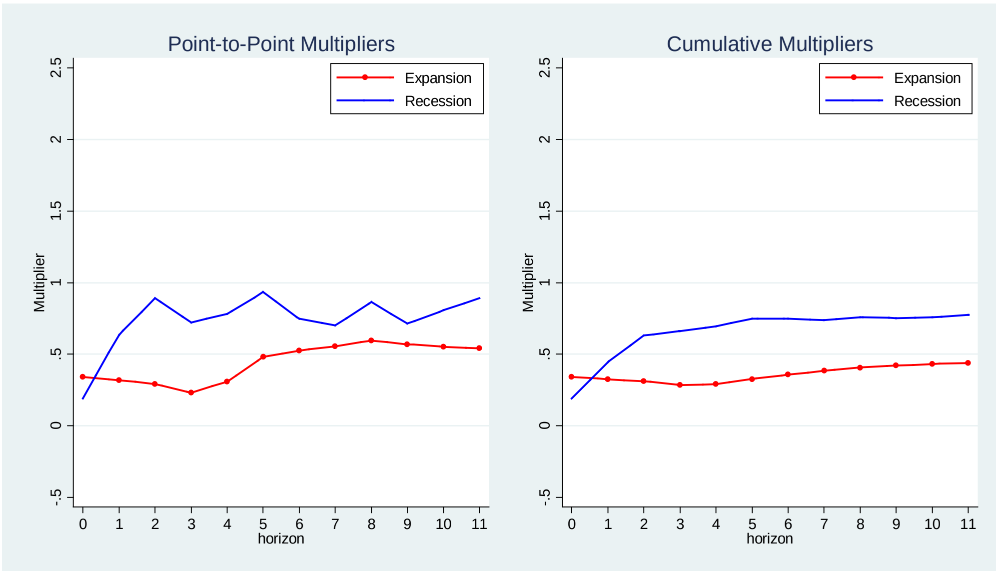
We will now show that the change in the variable being orthogonalized is the sole source of Yuriy's higher multipliers. To show this, we use all aspects of Yuriy's specification, but use the shock to $\Delta\ ln(g_t)$, orthogonalized with respect to four lags of $\Delta\ ln(g_t)$ and $\Delta\ ln(y_t)$, and a quartic time trend since this is the most similar to what Blanchard-Perotti did. Figure 4 below shows the effects of this one change. The results show that even Yuriy's point-to-point "multipliers" lie below one, and that the cumulative multipliers look very similar to those we reported in our Blanchard-Perotti identification robustness check. The multipliers in recession are slightly above those in expansion but are around 0.75, similar to what we report for our robustness check. Since we do not advocate the Blanchard-Perotti identification scheme, we do not take a stand on which implementation of their scheme is preferable.

Figure 4. Yuriy's specification, but with shock to growth of government spending



Suppose, however, one prefers Yuriy's method of using the Hall-Barro-Redlick-type transformation on the instruments, but for consistency extends that to the control variables as well. In particular, suppose that one uses Yuriy's variable for government spending shown above (with potential GDP in the denominator), and includes as control four lags of that variable as well as the analogous variable for real GDP (instead of the standard log variables). Figure 5 shows these results:

Figure 5. Yuriy's Specification, but with his version of the Hall-Barro-Redlick transformation as control variables.



Thus, Yuriy's finding of recession multipliers that are substantially above one in recessions disappears once one uses standard ways of calculating multipliers or changes small details in the implementation of the Blanchard-Perotti identification.

(2) First-Stage F-statistics

Yuriy showed graphs of the first-stage F-statistics for our military news variable by horizon. Yuriy's F-statistic is the square of the t-statistic on the coefficient of the shock in a regression of the government spending variable on the shock plus controls. This regression is the first-stage for Yuriy's point-to-point multipliers, and is not directly informative about cumulative multipliers. Nevertheless, we agree that these F-statistics contain useful information and are probably a better way of indicating relevance than the joint test we used for horizon 0. Yuriy argued that the F-statistics for military news during the high unemployment state become smaller for samples that exclude World War II. This last statement is not something we contest; indeed, our argument for including historical data with wars is that they add important variation.

What Yuriy didn't show was the same F-statistic across horizons for the leading alternative shock - the Blanchard-Perotti shock. The following graphs show first-stage F-statistics for both our military news variable and the standard Blanchard-Perotti shock. Figure 6 shows the F-statistics for the full sample, for the linear model and for the state-dependent model. We draw a line at 10 since that is a rough "safety" threshold for instrument relevance. (Most implicit first-stage F-statistics in macro are unfortunately well below 10.) All of the graphs make clear that while the Blanchard-Perotti shock has higher F-statistics for the first few horizons, they are very low after the first few quarters. In contrast, the military news shock does better as the horizon increases. This is to be expected, because as Ramey (2011) argued, the news variable is about *future* government spending and hence does not translate into an immediate increase in government spending. However, it does translate into an immediate jump in output because agents act on the news. (The comparable F-statistic for the output variable at horizon 0 is 14.6.) In contrast, the Blanchard-Perotti shock has an F-statistic above 5 for only the first three quarters in the high unemployment state for the full sample.

Figure 6. F-statistics for Identifying Yuriy's Point-to-Point Multipliers: Full Sample (Note that F-statistics are capped at 20)

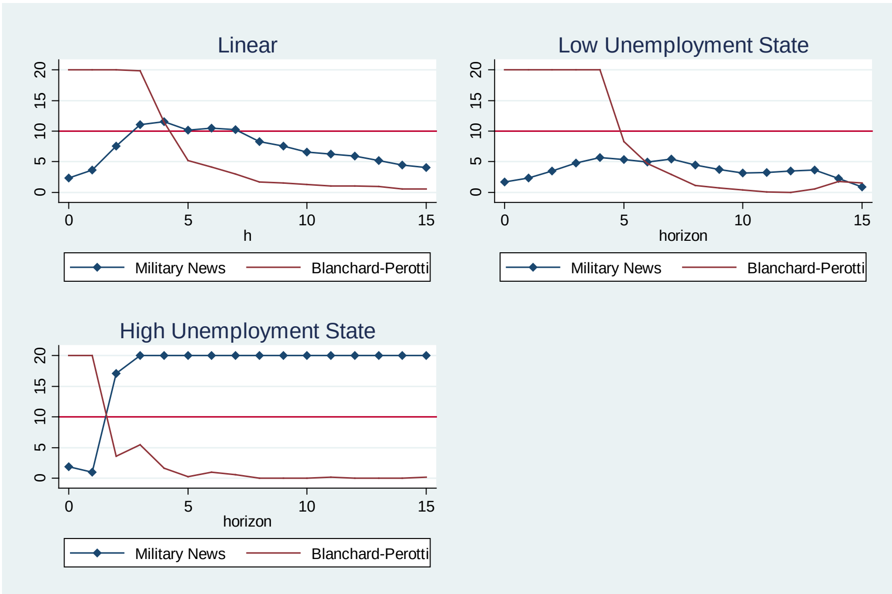


Figure 7 shows that when dependent variable observations that occur during WWII quantity constraints (defined as the second quarter of 1941 through the end of 1945) are omitted, the F-statistics for both shocks fall at most horizons, and particularly in the high unemployment state. Thus, these results support our contention that there is much less information in the data if we exclude WWII.

Figure 7. F-statistics for Identifying Yuriy's Point-to-Point Multipliers, Excluding WWII

(Note that F-statistics are capped at 20)

References

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

Auerbach, Alan and Yuriy Gorodnichenko, "Fiscal Multipliers in Recession and Expansion," in *Fiscal Policy After the Financial Crisis,* edited by Alberto Alesina and Francesco Giavazzi, pp. 63-98, 2013, 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.

Fisher, Jonas D.M., and Ryan Peters, "Using Stock Returns to Identify Government Spending Shocks," *The Economic Journal*, 120 (May 2010): 414-436.

Mountford, Andrew, and Harald Uhlig, "What are the Effects of Fiscal Policy Shocks?" *Journal of Applied Econometrics*, 24 (April 2009): 960-992.

Ramey, Valerie A., "Identifying Government Spending Shocks: It's All in the Timing," *Quarterly Journal of Economics*, 126 (February 2011): 1-50.

Uhlig, Harald, "Understanding the Impact of Fiscal Policy: Some Fiscal Calculus," *American Economic Review Papers and Proceedings* 100 (May 2010): 30-34.