

Supplementary Appendix to Government Spending  
Multipliers in Good Times and in Bad: Evidence  
from U.S. Historical Data

Valerie A. Ramey

University of California, San Diego and NBER

Sarah Zubairy

Texas A&M University

9 May 2017

## Contents

<b>1</b>	<b>Biases in computing multipliers</b>	<b>4</b>
<b>2</b>	<b>Robustness of slack estimates</b>	<b>4</b>
<b>3</b>	<b>Robustness of ZLB estimates</b>	<b>6</b>
<b>4</b>	<b>Behavior of taxes</b>	<b>7</b>
<b>5</b>	<b>Comparison with Auerbach-Gorodnichenko (2012)</b>	<b>9</b>
<b>6</b>	<b>Comparison with Auerbach-Gorodnichenko (2013)</b>	<b>13</b>
<b>7</b>	<b>Comparison with Fazzari, Morley and Panovska (2015)</b>	<b>13</b>

## List of Tables

1	Estimates of Multipliers Across States of Slack: Accounting for Present Value Discounting . . . . .	16
2	Robustness check I: Estimates of Multipliers Across States of Slack . . . . .	17
3	Robustness check II: Estimates of Multipliers Across States of Slack . . . . .	18
4	Robustness checks I: Estimates of Multipliers Across Monetary Policy Regimes	19
5	Robustness checks II: Estimates of Multipliers Across Monetary Policy Regimes	20
6	Estimates of Multipliers Across Monetary Policy Regimes: Excluding World War II and controlling for taxes and inflation . . . . .	21
7	Comparison to Auerbach-Gorodnichenko (2012) Multipliers . . . . .	22
8	Fazzari et al. (2015) Multipliers using the Jorda Method . . . . .	23

## List of Figures

1	Evolution of variables during war episodes . . . . .	24
2	Government spending and GDP responses to a Blanchard-Perotti shock: Considering slack states . . . . .	25
3	Cumulative multipliers to a Blanchard-Perotti shock: Considering slack states	26

4	Robustness check:Smooth transition threshold based on moving average of output growth . . . . .	27
5	Taxes and Deficit Responses to a news shock: Considering linear model . .	28
6	State-dependent Taxes and Deficit Responses to a news shock: Considering slack states . . . . .	29
7	Government spending and GDP responses to a Blanchard-Perotti shock: Considering zero lower bound . . . . .	30
8	Cumulative multipliers with Blanchard-Perotti shock: Considering zero lower bound . . . . .	31
9	Government spending and GDP responses to a news shock: Considering zero lower bound and excluding World War II . . . . .	32
10	Cumulative multipliers with news shock: Considering zero lower bound and excluding WWII . . . . .	33
11	Cumulative multipliers with Blanchard-Perotti shock: Considering zero lower bound and excluding WWII . . . . .	34
12	Estimating Auerbach and Gorodnichenko (2012) State Dependent Responses with the Jordà method . . . . .	35
13	Government spending and GDP responses to a news shock: Considering ZLB and Recessions in a Threshold-VAR . . . . .	36
14	Government spending and GDP responses to a Blanchard-Perotti shock: Considering ZLB and Recessions in a Threshold-VAR . . . . .	37

# 1 Biases in computing multipliers

To determine whether using *ex post* conversion factors can lead to inflated multipliers, we conducted a test based on the following point made by Ramey (2013). If the contemporaneous multiplier exceeds one, then it must be the case that private spending  $Y - G$  rises when  $G$  rises. Thus, one can compare the multipliers estimated the standard way to the response of real private spending to see if there is a contradiction.

To conduct this test, we first estimate a trivariate SVAR with military news, log real per capita government spending, and log real GDP, using four lags and quartic trend, on data from 1889 - 2015. The estimated elasticity at 2 year horizon is around 0.17 (based on the ratio of response of  $\ln(Y)$  to response of  $\ln(G)$  at 2 year horizon). We then multiply the estimated elasticity by the average of  $Y/G$  for the full sample, and obtain an implied multiplier of 1.32. To conduct the comparison, we next estimate a model in which we substitute the log of real private spending for log real GDP, and compute the impulse response functions (using the standard method). These responses show that private spending *falls* when government spending rises, and specifically has a negative response at the 2 year horizon. Thus, these results imply a multiplier that is less than unity. It appears that the practice of backing out multipliers using *ex post* conversion factors can lead to upward biased multiplier estimates in some situations.

# 2 Robustness of slack estimates

Figure 2 shows the impulse responses using the Blanchard-Perotti shock. Figure 3 shows the cumulative multipliers based on the Blanchard-Perotti shock for all horizons up to 20 quarters.

We also conducted various additional robustness checks for both shocks. First, we compute multipliers using present values rather than simple sums. Table 1 shows that the results are very similar.

Second, we show the results for various transformations of the data. First, we follow Hall (2009) and Barro and Redlick (2011) in converting GDP and government spending changes to the same units before the estimation. In particular, our output and government spending variables on the left-hand-side are defined as  $(Y_{t+h} - Y_{t-1})/Y_{t-1}$  and  $(G_{t+h} - G_{t-1})/Y_{t-1}$ . The

first variable can be rewritten as:

$$\frac{Y_{t+h} - Y_{t-1}}{Y_{t-1}} \approx (\ln Y_{t+h} - \ln Y_{t-1})$$

and hence is analogous to the standard VAR specification. The second variable can be rewritten as:

$$\frac{G_{t+h} - G_{t-1}}{Y_{t-1}} \approx (\ln G_{t+h} - \ln G_{t-1}) \cdot \frac{G_{t-1}}{Y_{t-1}}$$

Thus, this variable converts the percent changes to dollar changes using the value of  $G/Y$  at each point in time, rather than using sample averages. This means that the coefficients from the  $Y$  equations are in the same units as those from the  $G$  equations, which is required for constructing multipliers. With this specification, we use a quartic trend and lags of logged GDP and logged government spending on the right hand side. Table 2 shows that our baseline results still hold. For the case of Blanchard and Perotti shocks, the multipliers are larger in magnitude in the linear and state dependent case but there is no evidence of higher multipliers in high unemployment state.

The second panel of Table 2 also shows the results where in our baseline specification, we use an alternative definition of trend GDP. We construct potential GDP using a cubic trend for the early sample, fitted excluding Great Depression and append with the CBO potential GDP measure for the available years. This alternative measure of potential GDP yields very similar multipliers to our baseline.

Also, we investigated the impact of using a different interpolation method for the data. Recall that our underlying interpolators were quite volatile and led to volatility in the early data. To investigate the impact, we create alternative data that uses linear interpolation of the annual data in the pre-WWII period. The top panel of Table 3 shows that the results for the military news shock are little changed. Those for Blanchard-Perotti shock show slightly higher multipliers (0.5 to 0.6 in the linear case) and estimates that are virtually the same across states. The four last panels of Table 3 show various other explorations, such as including trends, using Christina Romer's estimates of historical GDP and unemployment, controlling for taxes and inflation, and controlling for trade-to-GDP ratios. The results are similar to the baseline.

### 3 Robustness of ZLB estimates

Tables 4 and 5 show various robustness checks for the ZLB state dependence case. When we consider Hall-Barro-Redlick transformation for GDP and government spending, the multipliers are higher overall, particularly for Blanchard and Perotti shocks, but there is no evidence of state-dependence. Using an alternative measure of trend GDP based on CBO potential GDP measure yields results very similar to the baseline. When we consider linearly interpolated data, the estimates for the ZLB state are slightly lower than the baseline case, and higher at lower horizons in the normal state. For the Blanchard and Perotti shock the estimated multipliers are also higher in the normal state. The last panel shows that the results are also similar to baseline for military shocks when we control for trade-to-GDP ratio. However, the multipliers for the Blanchard and Perotti shock are larger than the baseline in the normal case. While the addition of trends does not affect the military news shock results, the multipliers for the Blanchard and Perotti shock are larger than the baseline in the normal state, in this case too. In fact, with the inclusion of trends, the multiplier in the normal state becomes as large as the ZLB state at the 4 year horizon. In addition, Table 5 and Table 6 show multiplier estimates for both the full sample and excluding WWII when taxes and inflation are added as controls. The controls raise the multiplier estimates for the two-year horizon for the case of excluding WWII to 1.8, but the standard error is also higher and the multipliers are not statistically different across states.

Figure 9 shows the impulse responses for the baseline ZLB and normal states estimated on the sample excluding WWII using military news shock. In contrast to the full sample (shown in the text), where government spending rises robustly beginning two quarters after the news shock in the ZLB state, in the sample excluding WWII government spending becomes slightly negative for two quarters then slightly positive for two more quarters before beginning to rise robustly at horizon 5 with a peak at horizon 7. Output, however, rises steadily as soon as the news hits and then jumps even higher at horizon 5. The delay in government spending after the news, along with the anticipation effect in GDP, shows up in the multipliers in Figure 10. The estimated ZLB multipliers swing wildly from large negative to large positive with very wide confidence bands for the first few quarters, then reach 1.44 by horizon 8. Recall from the text that the military news instrument rises above the relevance threshold at horizon 5. At horizon 5, the multiplier is estimated to be 2.18 with a standard error of 0.33. However, the IRFs demonstrate that the size of the multiplier stems from output responding more quickly to the news than government spending does.

## 4 Behavior of taxes

Our analysis did not explicitly study the responses of taxes. Table 3 shows that the multipliers do not change when we control for taxes. However, the multipliers reported are based on the average response of taxes in the sample. Romer and Romer's (2010) estimates of tax effects indicate very significant negative multipliers on taxes, on the order of -2 to -3. Thus, it is important for us to consider how the increases in government spending are financed in order to interpret our multiplier results.

To analyze how taxes and deficits behave, we re-estimate our basic model augmented to include deficits and tax rates so that we can distinguish increases in revenues caused by rising output versus rising rates. Average tax rates are computed as the ratio of federal receipts to nominal GDP. The deficit is the real total deficit. We include four lags of these two new variables along with GDP and government spending as controls in all of the regressions, and we estimate this system for the full sample using the Jordà method.

Figure 5 shows the results from the linear case. The responses of government spending and GDP, as well as the multipliers, are almost identical to the baseline case. The middle panels show that both average tax rates and the deficit increase in response to news shock. Taking the ratio of the cumulative response of deficit to cumulative response of government spending at various horizons, we estimate a sharp rise in the share of government spending financed with the deficit during the first year. The deficit fraction of government spending then plateaus at 60 percent .

From a theoretical perspective, the fact that tax rates increase steadily during the first two years has significant implications for the multiplier. If all taxes are lump-sum taxes, news about a future increase in the present discounted value of government spending leads to an immediate jump in hours and output because of the negative wealth effect. In a neoclassical model, the effect is the same whether the taxes are levied concurrently or in the future. In contrast, the need to raise revenues through distortionary taxation can change incentives significantly. As Baxter and King (1993) show, if government spending is financed with current increases in tax rates, the multiplier can become negative in a neoclassical model.

The situation changes considerably when tax rates are slow to adjust, but agents anticipate higher future tax rates. To see this, consider the case of labor income taxes and a forward-looking household:

$$(1) \quad 1 = \beta E_t \left[ \frac{u_{n,t+1}}{u_{n,t}} \frac{(1 - \tau_t)w_t}{(1 - \tau_{t+1})w_{t+1}} (1 + r_t) \right]$$

where  $u_n$  is the marginal utility of leisure,  $\tau$  is the tax on labor income,  $w$  is the real wage rate,  $r$  is the real interest rate, and  $E_t$  is the expectation based on period  $t$  information. In expectation, the household should vary the growth rate of leisure inversely with the growth rate of after-tax real wages. This means that if  $\tau_{t+1}$  is expected to rise relative to  $\tau_t$ , households have an incentive to substitute their labor to the present (when it is taxed less) and their leisure to the future.

It is easy to show in a standard neoclassical model that the delayed response of taxes, such as we observe in the estimated impulse responses, results in a multiplier that is higher in the short-run but lower in the long-run relative to the lump-sum tax case. We have also conducted this experiment in the Gali et al. (2007) model where 50 percent of the households are rule-of-thumb consumers. We found the same effect in that model as well. Drautzburg and Uhlig (2015) analyze an extension of the Smets-Wouters model and also find that the timing of distortionary taxes is very important for the size of the multiplier. Given the impulse response of tax rates, and with these theoretical results in mind, it is very possible that our estimated multipliers are greater than we would expect if taxation were lump-sum.

Nevertheless, our finding that multipliers do not vary across states could be due to differential financing patterns. To determine whether this is the case, Figure 6 shows the state-dependent results in response to a news shock. As we showed before, both government spending and GDP rise more if a news shock hits during a slack state, even after adjusting the initial size of the shock. The bottom panels show that tax rates and deficits also rise more during recessions, but there are other interesting differences in the patterns. When we study the ratio of the cumulative deficit to cumulative government spending at each point in time along the path, we find that more of government spending is financed with deficits when a shock hits during a slack state.<sup>1</sup> For example, at quarter seven the ratio of the cumulative deficit to cumulative government spending is 61 percent if a shock hits during a slack state but only about 45 percent if the shock hits during a non-slack state. Thus, on average short-run government spending is financed more with deficits if the shock hits during a slack state. In addition, tax rates rise with a delay during the slack state relative to non-slack state. This would imply that the multiplier should be greater during times of slack. In fact, our estimates imply that it is not.

---

1. This is true with the exception of the second quarter. This can be explained by the fact that initially government spending and deficits rise slowly in response to a news shock and for the initial two quarters, deficits fall very slightly before rising.

## 5 Comparison with Auerbach-Gorodnichenko (2012)

This section discusses the details of the results we mentioned in the text in our discussion of Auerbach and Gorodnichenko's (2012) (AG-12) method. AG-12 use a smooth transition VAR (STVAR) model, post-WWII data, the Blanchard-Perotti identification method, and a function of the 7-quarter centered moving average of normalized real GDP growth as their measure of the state. They use three lags of the endogenous variables in the VAR, but also include four lags of the 7-quarter moving average growth rate as exogenous regressors in their model.<sup>2</sup> They construct their baseline impulse responses based on two assumptions: (1) the economy remains in an extreme recession or expansion state for at least 5 years; and (2) changes in government spending do not impact the state of the economy.<sup>3</sup> They find multipliers during recessions that are well above two and these results have been cited by those advocating stimulus during the Great Recession (e.g. DeLong and Summers (2012))

To understand the difference between our results and theirs, we begin by taking only one step away from what AG-12 did by using all details of their analysis except the method for estimating and constructing the impulse responses. In particular, we apply the Jordà method to their post-WWII data, using their exact definition of states, logs of variables, estimated government spending shocks from their STVAR model, and inclusion of four lagged values of the centered 7-quarter moving average of output growth as controls.  $F(z)$  is the indicator of the state as a function of the moving average of output ( $z$ ). It varies between a maximum of one (extreme recession) and zero (extreme expansion).

The top panel of Figure 12 reproduces the impulse responses shown in Figure 2 from AG-12. They construct these impulse responses from their STVAR estimates, assuming that the economy does not switch states for at least 20 quarters. The second panel shows the linear responses estimated using the Jordà method.<sup>4</sup> The government spending response looks similar to the linear case in AG-12, though the GDP response is more erratic and the standard error bands are much wider. The state-dependent responses shown in the lower panel look very different. The Jordà method produces impulse responses in which the response of government spending to a shock is higher in a recession than in an expansion, similar to our earlier results, but in opposition to those of AG-12. The response of output differs little across

---

2. The published paper does not discuss these additional terms, but the initial working paper version includes these terms in one equation and the codes posted for the published paper include them.

3. These results are shown in Table 1 and Figure 2 of their paper.

4. We have multiplied the log output response by a conversion factor of 5.6, following AG-12. They use a nonstandard measure of government purchases as their measure of G. As a result their  $Y/G$  ratio used to convert multipliers is higher than the usual one based on total government purchases.

states, in contrast to AG-12 who find that output rises robustly and continuously throughout the 5 years in the recession state but quickly falls toward zero and becomes negative in the expansion state.

The first panel of Table 7 compares AG-12's cumulative 5-year and 2-year multipliers to those we estimated by the Jordà method. For the 5-year horizon, AG find multipliers of 2.24 for recessions and -0.33 for expansions whereas the Jordà method estimates multipliers of 0.84 in recessions and -0.59 in expansions.<sup>5</sup> Thus, the Jordà multipliers are below one in recessions, but similar to AG-12's multipliers in expansions. At the 2-year horizon, AG's estimates imply a recession multiplier of 1.65 and a gap between states of 1.55. In contrast, the Jordà method implies a 2-year multiplier in recessions of 0.24 and a gap between states of -0.12.<sup>6</sup>

Thus, even when we use the same sample period, data, variable definitions, definition of slack, and estimated shocks as AG-12, the Jordà method produces multipliers in recessions that are much lower than those of AG-12. This means that AG-12's high multipliers during recessions are likely due to the method for constructing the impulse responses. To see the importance of these assumptions, we conduct several experiments. In these experiments, we compute alternative impulse responses by iterating on AG's STVAR parameter estimates under different assumptions about the dynamics of the state of the economy. Since the economy is never literally in an extreme recession or expansion, we focus on the average of "severe" recessions and "severe" expansions, which we define as the few quarters in which  $F(z)$  is above 0.95 or below 0.05, respectively. The few quarters of severe recession occur during the recessions of 1974-75 (two quarters), 1981-82 (one quarter), and 2008-09 (five quarters).

The second panel of Table 7 reports these experiments. For reference, the first line of the second panel shows that AG's baseline 5-year multipliers do not change much when we change  $F(z)$  by a small amount. The second line shows the multiplier calculated assuming a constant state and no feedback, but looking at the 2-year integral multiplier. This calculation requires less drastic assumptions because it only assumes that the state remains constant for 2 years rather than 5 years. Here, the multipliers in severe recessions are not as high and those in severe expansions are not as low, so the difference across states falls from 2.47 to 1.46.

The next experiment, "Actual State Dynamics," assumes that instead of staying constant

---

5. AG-12 also report a recession multiplier of 2.5, but that is based on comparing the peak response of output to the initial government spending shock, a practice we critiqued in a previous section.

6. One should keep in mind, however, that the Jordà estimates are not very precise. We were not able to estimate the standard errors of the multipliers because the `xtscc` command in Stata reported that the variance matrix was nonsymmetric or highly singular.

at an extreme value, the state indicator  $F(z)$  is equal to its actual value at each point in time. In practice the experiment is conducted as follows. We first calculate the paths of government spending and output, assuming that the shocks to the government spending, tax, and output equations take their estimated values. This essentially reproduces the actual path of the economy for all variables including  $F(z)$ . We then calculate an alternative path of government spending, taxes, and output assuming that there is an additional one-time positive shock in the current period to government spending, equal to one standard deviation of the estimated government spending shock (equal to 1.3 percent of government spending). We allow the shock to change the path of spending, taxes, and output, but not the state of the economy,  $F(z)$ , relative to its actual path. The difference between this simulation and the actual values of the variables forms the impulse response functions. Despite the lack of feedback, this experiment is different from AG's baseline experiment because it allows the state of the economy to experience its natural dynamics (i.e.  $F(z)$  is allowed to vary between its extremes as it actually does).<sup>7</sup> The third and fourth lines of the lower panel show the multipliers for this experiment. In severe recessions, the 5-year multiplier is estimated to be 1.4 and the 2-year multiplier is estimated to be 1.1. Thus, allowing the state of the economy to follow its subsequent natural dynamics reduces the constructed multiplier in recessions. The effect is not so big on the expansion multipliers, however. This is to be expected since expansions have a much longer duration than recessions, so the assumption of no change in state is not so at odds with the data.

AG-12 relax one of their baseline assumptions in a second experiment by allowing partial feedback of government spending on the state of the economy, but otherwise not allowing the state to change. They are not able to allow full feedback, though, because of the nature of their state variable.<sup>8</sup> The fifth and sixth lines of the second panel of the table show the average of their multipliers in severe recessions and expansions. Their experiment also lowers the estimated multiplier in severe recessions, to 1.36 for the 5-year multiplier and 1.01 for the 2-year multiplier.

The final two lines of the table show our experiment in which we allow both the natural dynamics of the economy and partial feedback from the government spending shock. The

---

7. The assumptions regarding the state  $F(z)$  are important. For instance, Caggiano et al. (2015) employ the STVAR approach of AG-12 for a shorter sample, but compute impulse response functions using the generalized impulse response approach advocated by Koop et al. (1996), and find that the spending multipliers in recessions are not statistically larger than in expansions. They only find evidence of nonlinearities when focusing on deep recessions versus strong expansionary periods.

8. Their state variable is a function of a *centered* moving average of GDP growth. Thus, future values of GDP enter the current state. This formulation makes it impossible to allow full feedback in a logically consistent manner.

experiment is the same as the "Actual State Dynamics, No Feedback" except that it also allows the  $F(z)$  indicator to change from its actual path in reaction to current changes of output resulting from the government spending shock. As shown in the table, both the 5-year and 2-year multipliers during severe recessions are calculated to be 1.07. During severe expansions, the 5-year multiplier is calculated to be 0.14, which is higher than AG's multiplier of -0.3. As a result, the gap in multipliers across states shrinks.

Thus, even when we use AG's STVAR parameter estimates, we can get very different estimates for the multiplier. Differing assumptions made about transitions between states and the feedback of government spending to the state lead to very different estimates of multipliers. When we compute multipliers allowing for the natural dynamics of the economy we find a much smaller gap across states than AG-12. The gap we do find is not because the multiplier is so high during recessions, but because it is estimated to be so low during expansions when we use AG's data and variable definitions.

As pointed out above, the state variable that AG-12 employ is a function of a *centered* moving average of GDP growth. This suggest that future values of GDP are used to construct the current state of the economy. This formulation not only makes it impossible to allow full feedback in a logically consistent manner, but it in fact plays an important role in driving their results as well. In our exploration of the AG-12 results, we found that if we focus only on the backward moving average terms, and in particular, instead of a 7-quarter centered moving average growth rate, consider a 4-quarter backward moving average growth rate, while keeping all the details of the AG set-up the same, the resulting multiplier is much larger in expansion than in recession.<sup>9</sup> Alloza (2014) conducts a much more systematic analysis of the importance of this two-sided moving average filter and further corroborates our finding. He finds that using a higher order of centered moving average (i.e. using more information about future values of GDP) reduces the effect of the spending shock on output during times of booms and amplifies it during recession periods. He also explores varying the symmetry of the 7-quarter moving average, and shows that when less information about the future is used, then the results suggest a more robust response of output to a spending shock in booms, while the response in recessions becomes negative. This suggests, that the future information used in constructing the current state variable in AG-12 is also crucial to their results.

---

9. The 5 year cumulative multiplier is close to 0.7 in recession and 1.8 in expansion.

## 6 Comparison with Auerbach-Gorodnichenko (2013)

Auerbach-Gorodnichenko (AG-13) also applies the Jordà method to a panel of OECD countries; in fact, AG-13 were the ones who first realized the potential of this method for state-dependent fiscal models. Thus, a key question is why they find higher multipliers during recessions even with this method. There is, of course, the obvious difference in time period and country sample. We believe, however, the most likely reason for the difference is in two details of how they calculate multipliers. First, following the standard practice, they estimate everything in logarithms and then use the *ex post* conversion factor based on average  $Y/G$  during their sample to convert elasticities to multipliers. Second, they follow Blanchard and Perotti (2002) by comparing the path or peak of output to the *impact* of government spending rather than to the peak or integral of the path of government spending. This is a big difference because the effects of a shock to government usually build up for several quarters. As we argued in a previous section, this is not the type of multiplier that interests policy makers because it does not count the average cumulative cost of government spending associated with the path. If we used that same procedure on our baseline estimates, calculating multipliers by dividing the average response of output over the 2-year horizon by the *initial* shock to government spending, we would produce multipliers of 4.3 in the linear case, 2.2 in the low unemployment state and 8.8 in the high unemployment state! Thus, it is clear that even using the same estimation method and same method for computing impulse responses, details of the calculations of multipliers can make a big difference.

## 7 Comparison with Fazzari, Morley and Panovska (2015)

Fazzari et al. (2015) estimate a threshold structural VAR and extend Auerbach and Gorodnichenko's (2012) work in three key ways: (i) they estimate multipliers based on generalized impulse responses functions; (ii) they investigate a variety of measures of slack; and (iii) they estimate the threshold for each measure rather than imposing it *a priori*. Based on a series of statistical tests, they use their own structural break-adjusted capacity utilization index and estimate their model from 1967-2012. They find evidence of a higher multiplier, of 1.6, in the slack state, defined as a structural break-adjusted capacity utilization rate below their estimated threshold.

To understand the difference between our results and theirs, we begin by taking only one step away from what they did by using all details of their analysis except the method for

estimating and constructing the impulse responses. In particular, we employ their sample choice, identification scheme, specification of variables and threshold variables along with the estimated threshold values, but use the Jorda methodology to estimate the multipliers. We conduct this exercise for their baseline specification that uses their structural break-adjusted capacity utilization measure as the state variable; this measure subtracts a different mean for the 1967-1974 sample than for the post-1974 sample and this adjustment affects the estimated threshold for the entire sample. We also conduct this exercise for the other slack variables identified in their Table 2 that have marginal likelihoods close in size to the best fitting model.

Table 8 shows the implied multipliers. The first panel shows that when we use their preferred threshold variable of adjusted capacity utilization, we also estimate larger multipliers in the slack state. At the 2-year horizon, it is as large as 2.5 and statistically significantly different from the multiplier in the high capacity utilization state. Thus, our data and method produce even higher multipliers than they find. But then we consider their other slackness measures, i.e. the standard series on capacity utilization (without their 1967-1974 mean adjustment), the CBO output gap, and the unemployment rate as slack variables, using their threshold estimates. These results show that we no longer find any evidence of higher multipliers in the slack state, and in fact, in the case of unadjusted capacity utilization, the multipliers are *negative* in the slack state. The only case where we find higher multipliers in slack is the case of the unemployment rate threshold, but the multipliers are much smaller than one under both states, not statistically significantly different across states and the point estimate difference in the long run is driven by multipliers being low in the non-slack state. These results suggest that the multipliers produced by the Fazzari et al. (2015) method are not robust to closely-related slack measures and are highly dependent on the specific choice of threshold variable and threshold estimate. Multipliers estimated using these close alternatives to the slack index are not shown in their paper.

Thus, the only one of their high marginal likelihood slack measures that produces high multipliers in recessions is the capacity utilization series that they adjust for a structural break in 1974q1. This adjustment, along with their estimated threshold, reassigns 20 percent of the sample to a different state relative to the standard capacity utilization index and threshold. We investigated several aspects of this series. First, we found that their structural break-adjusted capacity utilization index has a lower correlation with the CBO output gap than the unadjusted capacity utilization index (0.6 vs. 0.8). Second, when we tested their series for a single structural break, we found a break at 2001q1, not 1974q1. This break date held in tests

that trimmed the standard 15 percent of each end of the sample (which excludes early 1974) as well as in those that trimmed 5 and 10 percent. It also held for both the supremum Wald test and the supremum likelihood ratio test. We also tested a series on capacity utilization in manufacturing, since that series looks very similar to the total series but has the advantage that it is available back to 1948. A structural break test on that series found a break at 2000q4. These auxiliary tests raise questions about the structural break-adjusted series used by Fazzari et al. (2015), and hence the estimated multipliers associated with that series.

**Table 1. Estimates of Multipliers Across States of Slack: Accounting for Present Value Discounting**

	Linear Model	High Unemployment	Low Unemployment
<b>Baseline</b>			
<b>Military news shocks</b>			
2 year integral	0.70	0.43	0.69
4 year integral	0.71	0.56	0.60
<b>Blanchard-Perotti shocks</b>			
2 year integral	0.53	0.74	0.45
4 year integral	0.60	0.81	0.42
<b>Present discounted value multipliers</b>			
<b>Military news shocks</b>			
2 year integral	0.71	0.40	0.71
4 year integral	0.71	0.59	0.57
<b>Blanchard-Perotti shocks</b>			
2 year integral	0.51	0.73	0.44
4 year integral	0.60	0.80	0.42

Note: The multipliers are computed for the sub-sample 1920-2015, where the discount factor is constructed using the T-bill rate and past inflation.

**Table 2. Robustness check I: Estimates of Multipliers Across States of Slack**

	Linear Model	High Unemployment	Low Unemployment
Hall-Barro-Redlick transformation			
<b>Military news shock</b>			
2 year integral	0.76	0.75	0.82
4 year integral	0.83	0.79	0.97
<b>Blanchard-Perotti shock</b>			
2 year integral	0.57	0.69	0.48
4 year integral	0.93	0.82	1.02
Alternative definition of potential GDP based on CBO measure			
<b>Military news shock</b>			
2 year integral	0.67	0.61	0.64
4 year integral	0.71	0.69	0.68
<b>Blanchard-Perotti shock</b>			
2 year integral	0.39	0.70	0.32 <sup>†</sup>
4 year integral	0.48	0.79	0.35 <sup>†</sup>

Note: The symbol on the last entry in each row signifies the p-values for difference in multipliers across state, where † indicates HAC-robust p-value,  $p_{HAC} < 0.1$ .

**Table 3. Robustness check II: Estimates of Multipliers Across States of Slack**

	Linear Model	High Unemployment	Low Unemployment
Linearly interpolated data			
<b>Military news shock</b>			
2 year integral	0.66	0.67	0.53
4 year integral	0.69	0.72	0.57
<b>Blanchard-Perotti shock</b>			
2 year integral	0.51	0.42	0.47
4 year integral	0.58	0.57	0.53
Addition of trends			
<b>Military news shock</b>			
2 year integral	0.77	0.70	0.61
4 year integral	0.85	0.75	0.72
<b>Blanchard-Perotti shock</b>			
2 year integral	0.46	0.70	0.42 <sup>†</sup>
4 year integral	0.62	0.82	0.53 <sup>†</sup>
Christina Romer series for GDP, PGDP and UNEMP			
<b>Military news shock</b>			
2 year integral	0.66	0.49	0.61
4 year integral	0.69	0.62	0.61
<b>Blanchard-Perotti shock</b>			
2 year integral	0.40	0.67	0.33 <sup>†</sup>
4 year integral	0.48	0.723	0.34 <sup>†</sup>
Additional controls for taxes and inflation			
<b>Military news shock</b>			
2 year integral	0.67	0.66	0.56
4 year integral	0.71	0.69	0.60
<b>Blanchard-Perotti shock</b>			
2 year integral	0.38	0.67	0.37 <sup>†</sup>
4 year integral	0.44	0.79	0.40 <sup>†</sup>
Additional controls for trade-to-GDP ratio			
<b>Military news shock</b>			
2 year integral	0.66	0.60	0.59
4 year integral	0.71	0.68	0.67
<b>Blanchard-Perotti shock</b>			
2 year integral	0.38	0.68	0.30
4 year integral	0.47	0.77	0.35

Note: The symbol on the last entry in each row signifies the p-values for difference in multipliers across state, where <sup>†</sup> indicates HAC-robust p-value,  $p_{HAC} < 0.1$ .

**Table 4. Robustness checks I: Estimates of Multipliers Across Monetary Policy Regimes**

	Linear Model	Near Zero Lower Bound	Normal
Hall-Barro-Redlick transformation			
<b>Military news shock</b>			
2 year integral	0.76	0.78	0.54
4 year integral	0.83	0.74	0.74
<b>Blanchard-Perotti shock</b>			
2 year integral	0.57	0.69	0.54
4 year integral	0.93	0.81	1.09
Alternative definition of potential GDP based on CBO measure			
<b>Military news shock</b>			
2 year integral	0.67	0.78	0.56
4 year integral	0.71	0.77	0.59
<b>Blanchard-Perotti shock</b>			
2 year integral	0.39	0.66	0.06 <sup>†</sup>
4 year integral	0.48	0.74	0.06 <sup>†</sup>

Note: The symbol on the last entry in each row signifies the p-values for difference in multipliers across state, where <sup>†</sup> indicates HAC-robust p-value,  $p_{HAC} < 0.1$ .

**Table 5. Robustness checks II: Estimates of Multipliers Across Monetary Policy Regimes**

	Linear Model	Near Zero Lower Bound	Normal
Linearly interpolated data			
<b>Military news shock</b>			
2 year integral	0.66	0.66	0.37
4 year integral	0.69	0.69	1.03
<b>Blanchard-Perotti shock</b>			
2 year integral	0.51	0.61	0.31
4 year integral	0.58	0.69	0.35
Addition of trends			
<b>Military news shock</b>			
2 year integral	0.77	0.78	0.65
4 year integral	0.85	0.78	0.86
<b>Blanchard-Perotti shock</b>			
2 year integral	0.46	0.64	0.33
4 year integral	0.62	0.71	0.75
Additional controls for inflation			
<b>Military news shock</b>			
2 year integral	0.67	0.77	0.60
4 year integral	0.71	0.77	0.58
<b>Blanchard-Perotti shock</b>			
2 year integral	0.38	0.65	0.08 <sup>†</sup>
4 year integral	0.46	0.72	0.09 <sup>†</sup>
Additional controls for taxes and inflation			
<b>Military news shock</b>			
2 year integral	0.67	0.94	0.55
4 year integral	0.71	0.86	0.52
<b>Blanchard-Perotti shock</b>			
2 year integral	0.38	0.67	0.08 <sup>†</sup>
4 year integral	0.44	0.74	-0.02 <sup>†</sup>
Additional controls for trade-to-GDP ratio			
<b>Military news shock</b>			
2 year integral	0.66	0.76	0.63
4 year integral	0.71	0.75	0.78
<b>Blanchard-Perotti shock</b>			
2 year integral	0.38	0.64	0.10
4 year integral	0.47	0.70	0.12

Note: The symbol on the last entry in each row signifies the p-values for difference in multipliers across state, where † indicates HAC robust p-value  $p_{HAC} < 0.1$ .

**Table 6. Estimates of Multipliers Across Monetary Policy Regimes: Excluding World War II and controlling for taxes and inflation**

	Linear Model	Near Zero Lower Bound	Normal	P-value for difference in multipliers across
<b>Military news shock</b>				
2 year integral	0.80 (0.236)	1.73 (0.317)	0.55 (0.128)	HAC = 0.005 AR = 0.154
4 year integral	0.74 (0.170)	0.90 (0.159)	0.52 (0.224)	HAC = 0.154 AR = 0.174
<b>Blanchard-Perotti shock</b>				
2 year integral	0.06 (0.159)	0.85 (0.439)	0.08 (0.187)	HAC = 0.110 AR = 0.173
4 year integral	-0.07 (0.275)	-0.63 (1.57)	-0.02 (0.252)	HAC = 0.711 AR = 0.640

Note: The values in brackets under the multipliers give the standard errors. HAC indicates HAC-robust p-values and AR indicates weak instrument robust Anderson-Rubin p-values.

**Table 7. Comparison to Auerbach-Gorodnichenko (2012) Multipliers**

<b>Direct Comparisons</b>	Extreme Recession ( $F(z)=1$ )	Extreme Expansion ( $F(z)=0$ )	Difference
<b>AG's Estimates, Constant State</b>			
5 year integral	2.24 (0.24)	-0.33 (0.20)	2.57
2 year integral	1.65	0.10	1.55
<b>Jordà Method Applied to AG Specification</b>			
5 year integral	0.84	-0.59	1.43
2 year integral	0.24	0.36	-0.12
<b>Alternative Multipliers using AG's STVAR Estimates</b>	Severe Recession ( $F(z) \geq 0.95$ )	Severe Expansion ( $F(z) \leq 0.05$ )	Difference
<b>Constant State, No Feedback</b>			
5 year integral	2.16	-0.31	2.47
2 year integral	1.56	0.10	1.46
<b>Actual State Dynamics, No Feedback</b>			
5 year integral	1.41	0.19	1.22
2 year integral	1.13	0.15	0.97
<b>AG Partial Feedback</b>			
5 year integral	1.36	-0.04	1.40
2 year integral	1.01	0.15	0.86
<b>Actual State Dynamics, Partial Feedback</b>			
5 year integral	1.07	0.14	0.93
2 year integral	1.07	0.12	0.95

Note: STVAR denotes the Smooth Transition Vector Autoregression used by AG-12. Impulse responses are calculated based on the VAR parameter estimates and auxiliary assumptions. The values in brackets under the multipliers give the standard errors.  $F(z)$  is AG's indicator of the state of the economy.  $F(z) = 1$  indicates the most severe recession possible and  $F(z) = 0$  indicates the most extreme boom possible.

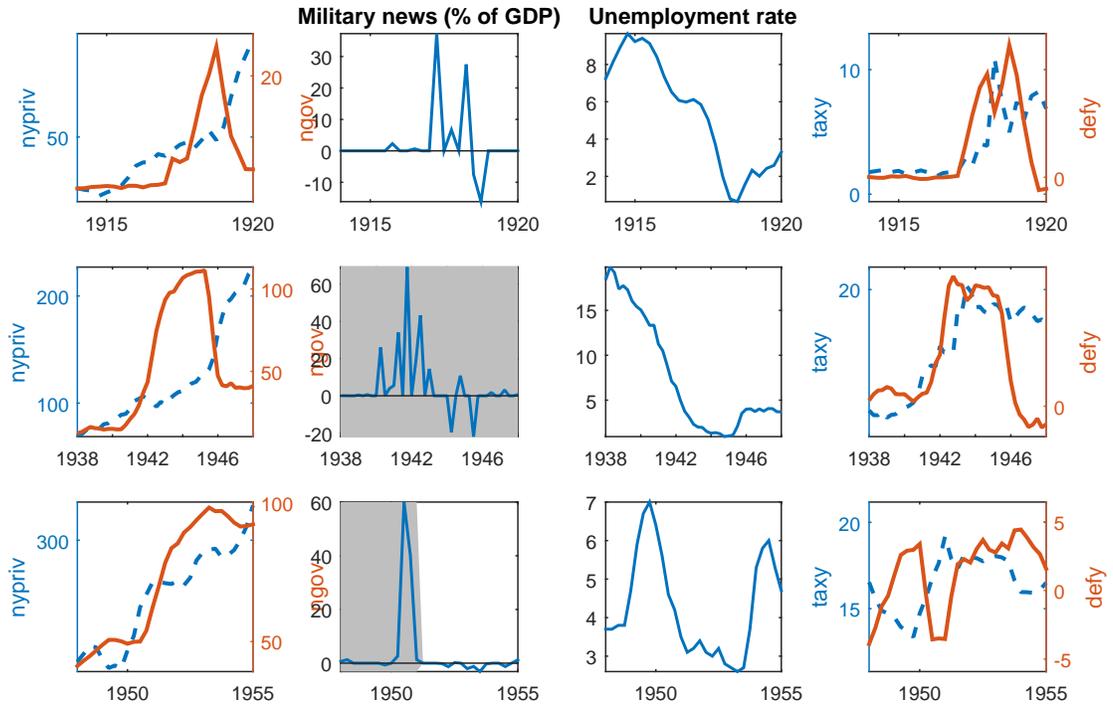
"Constant state" means that the impulse responses are calculated assuming that the economy remains in its current state for the duration of the multiplier. "Feedback" means that the estimates allow government spending to change the state of the economy going forward.

**Table 8. Fazzari et al. (2015) Multipliers using the Jorda Method**

	Linear Model	Slack State	Non-Slack State
Capacity utilization (adjusted)			
2 year integral	0.31	2.55	0.13 <sup>†</sup>
4 year integral	-0.50	1.33	-0.44
Capacity utilization			
2 year integral	0.31	-0.65	1.34
4 year integral	-0.50	-1.41	-0.32
Output gap			
2 year integral	0.31	0.69	0.96
4 year integral	-0.50	0.19	-0.31
Unemployment rate			
2 year integral	0.31	0.58	0.36
4 year integral	-0.50	0.13	-1.38

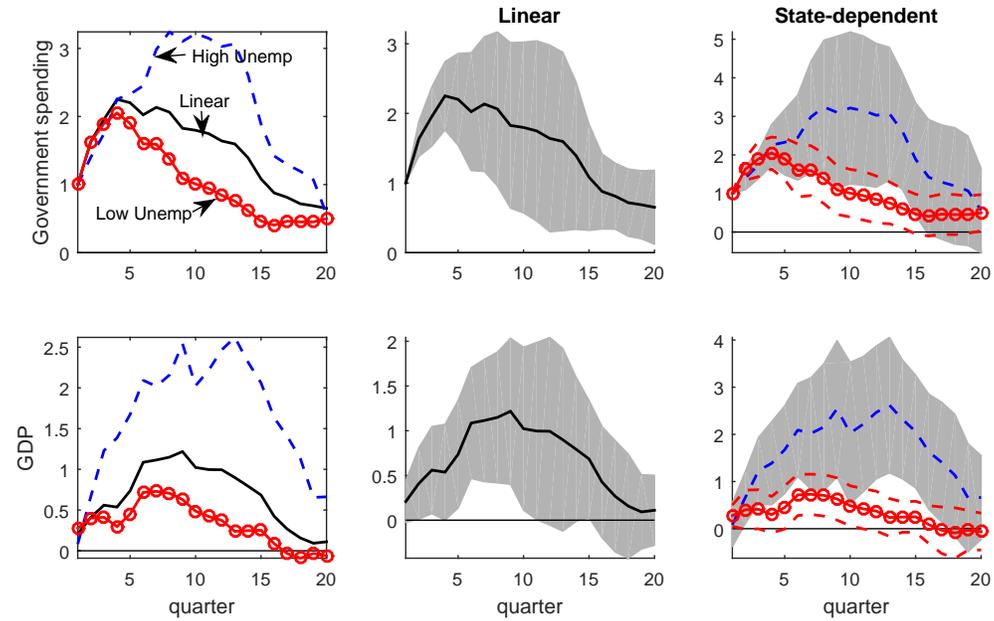
Note: We use all details of Fazzari et. al (2015) analysis including the sample choice, identification scheme, specification of variables, threshold variable along with the threshold estimate but use the Jorda methodology to get the multipliers. The symbol on the last entry in each row signifies the p-values for difference in multipliers across state, where † indicates HAC-robust p-value,  $p_{HAC} < 0.1$ .

**Figure 1. Evolution of variables during war episodes**



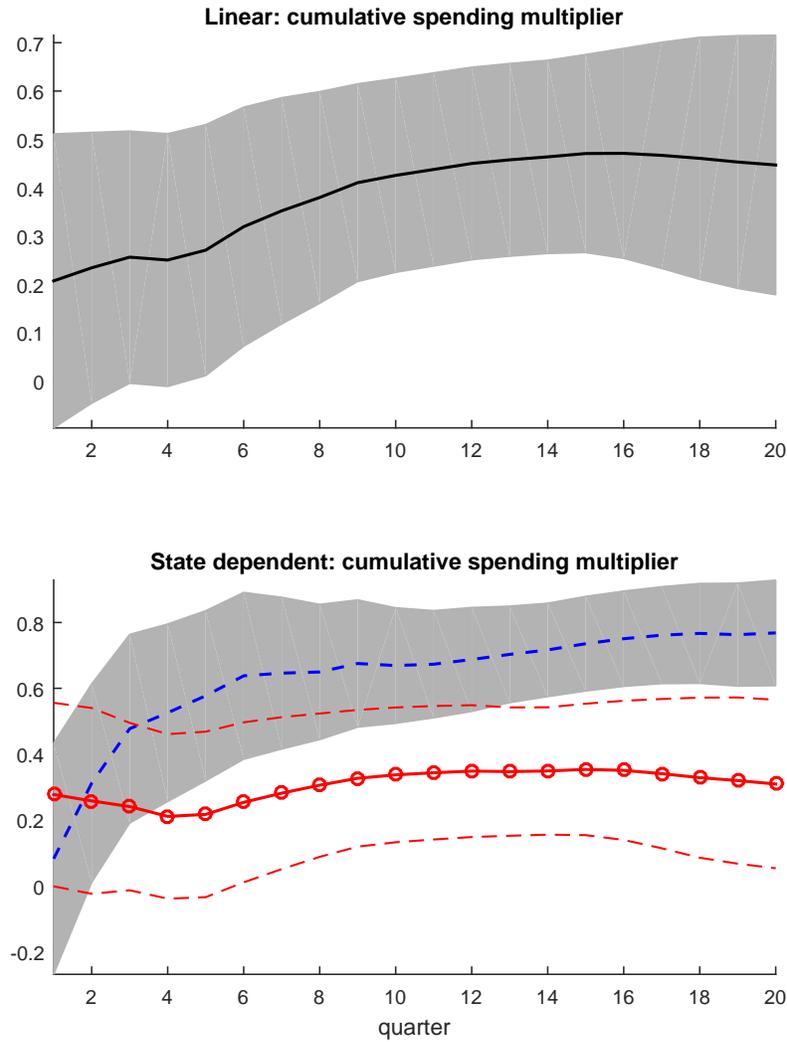
Note: The first column shows real private activity (left-axis) and real government spending (right axis). The second column shows military spending news with shaded areas indicating periods we classify as the zero lower bound period for interest rate, and the third column shows the civilian unemployment rate. The last column shows the average tax rate (left axis) and the deficit-to-GDP ratio (right axis). The first row corresponds to the period around World War I, the second row shows the time period around World War II and the last row shows the period around the Korean war.

**Figure 2. Government spending and GDP responses to a Blanchard-Perotti shock: Considering slack states**



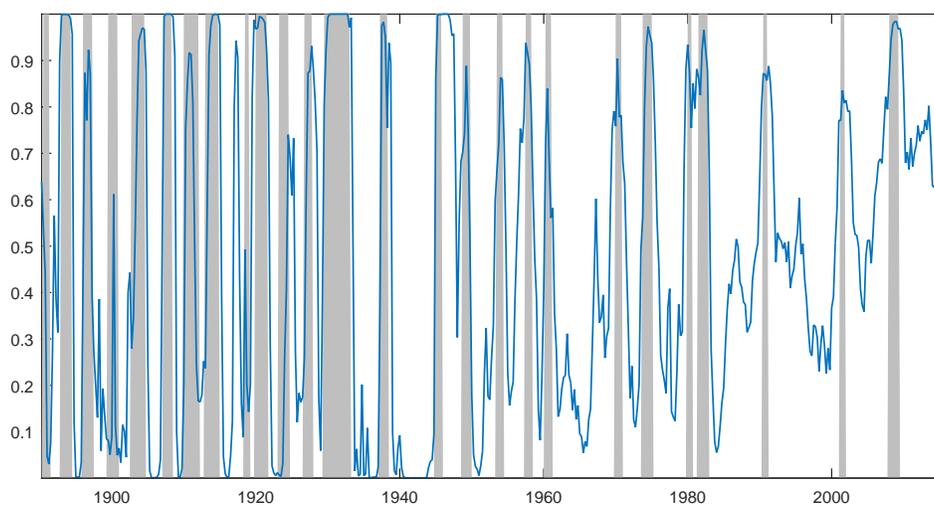
Note: Response of government spending and GDP to a 1% of government spending shock. The top row shows the response of government spending and the second row shows the response of GDP. The first column shows the responses in the linear and state-dependent model. The second column shows the responses in the linear model. The last column shows the state-dependent responses where the blue dashed lines are responses in the high unemployment state and the lines with red circles are responses in the low unemployment state. 95% confidence intervals are shown in second and third columns.

**Figure 3. Cumulative multipliers to a Blanchard-Perotti shock: Considering slack states**



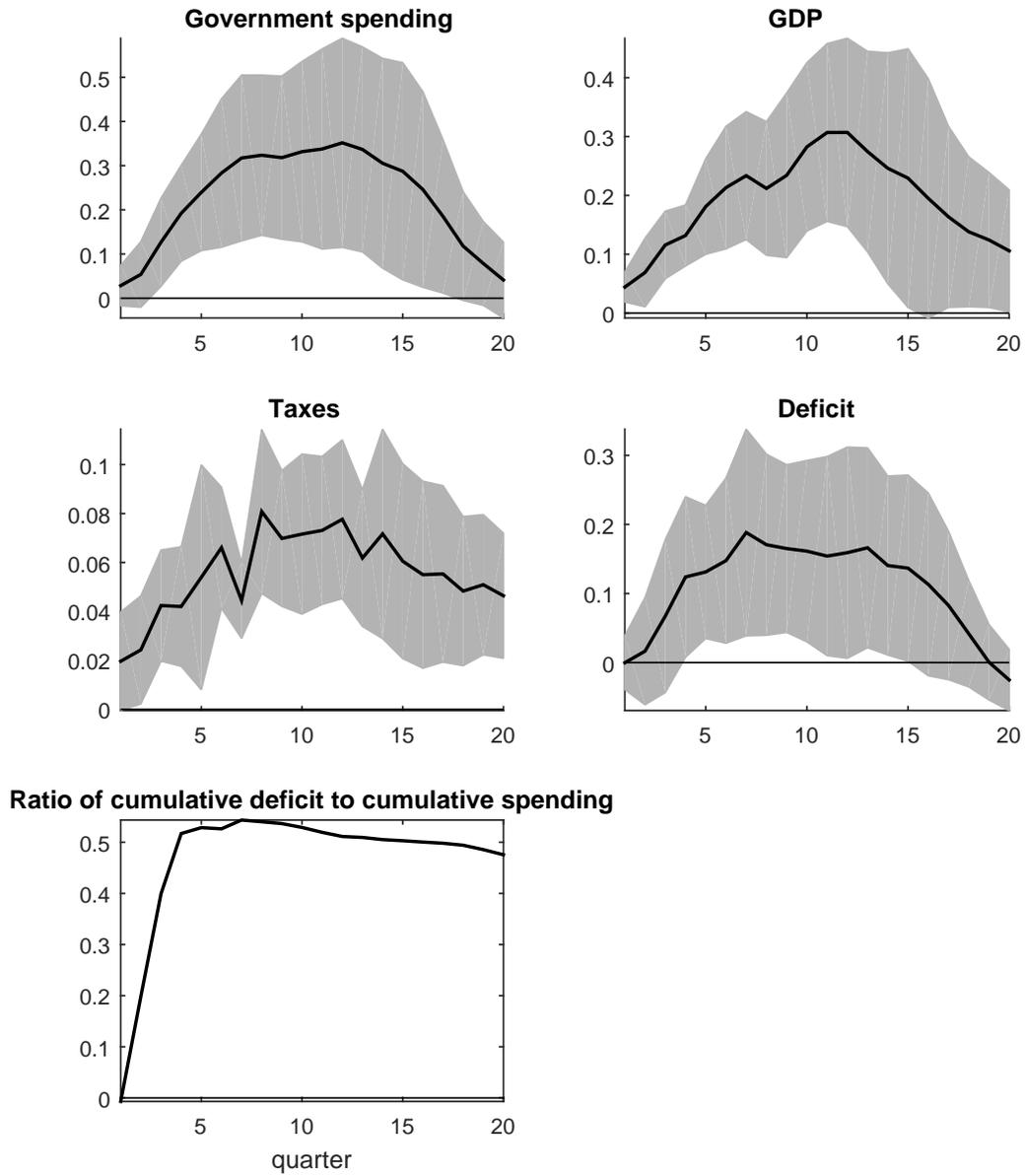
Note: Cumulative spending multipliers across different horizons. The top panel shows the cumulative multipliers in the linear model. The bottom panel shows the state-dependent multipliers where the blue dashed lines are multipliers in the high unemployment state and the lines with red circles are multipliers in the low unemployment state. 95% confidence intervals are shown in all cases.

**Figure 4. Robustness check: Smooth transition threshold based on moving average of output growth**



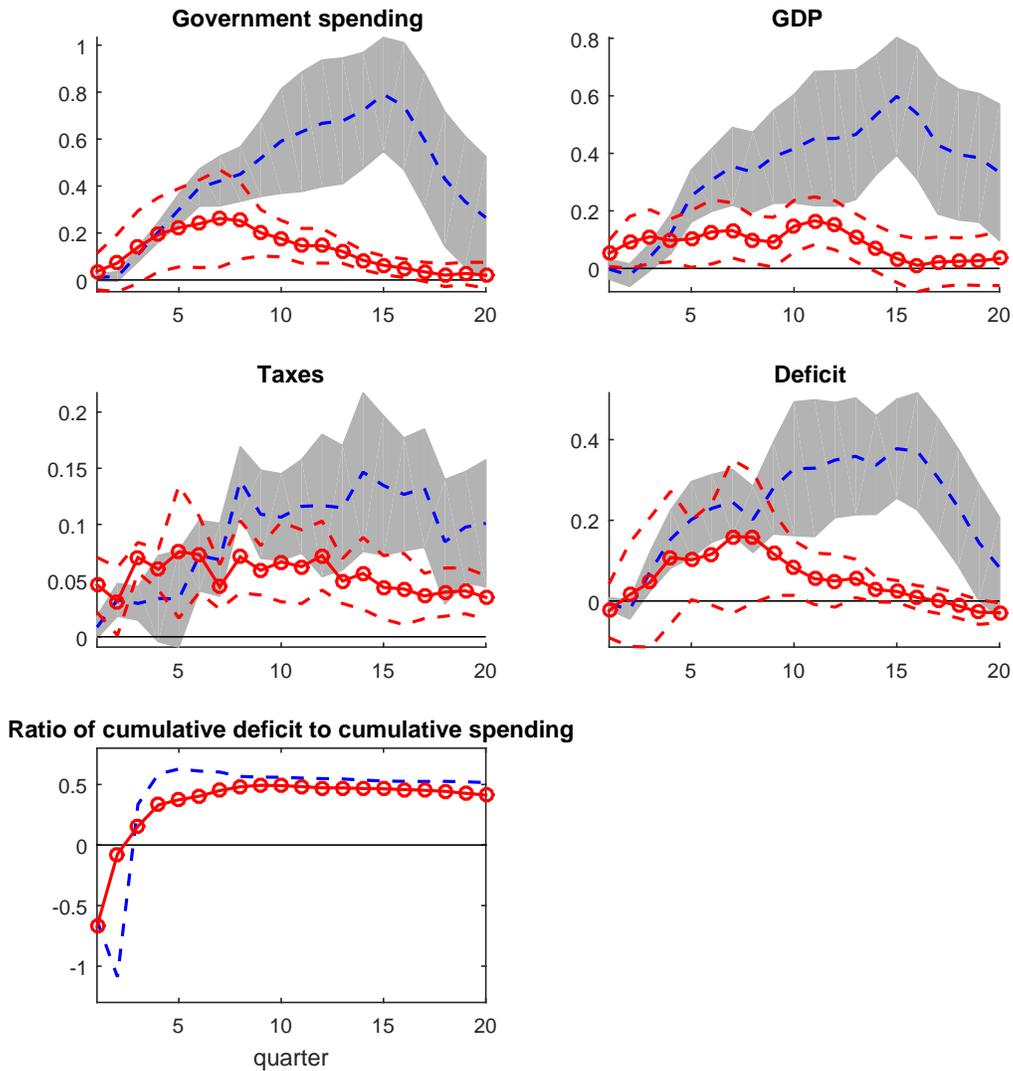
Note: The figures shows the weight on a recession regime, which is constructed using the 7 quarter centered moving average of output growth and the same definition as Auerbach and Gorodnichenko (2012). The shaded areas indicate NBER official recessions.

**Figure 5. Taxes and Deficit Responses to a news shock: Considering linear model**



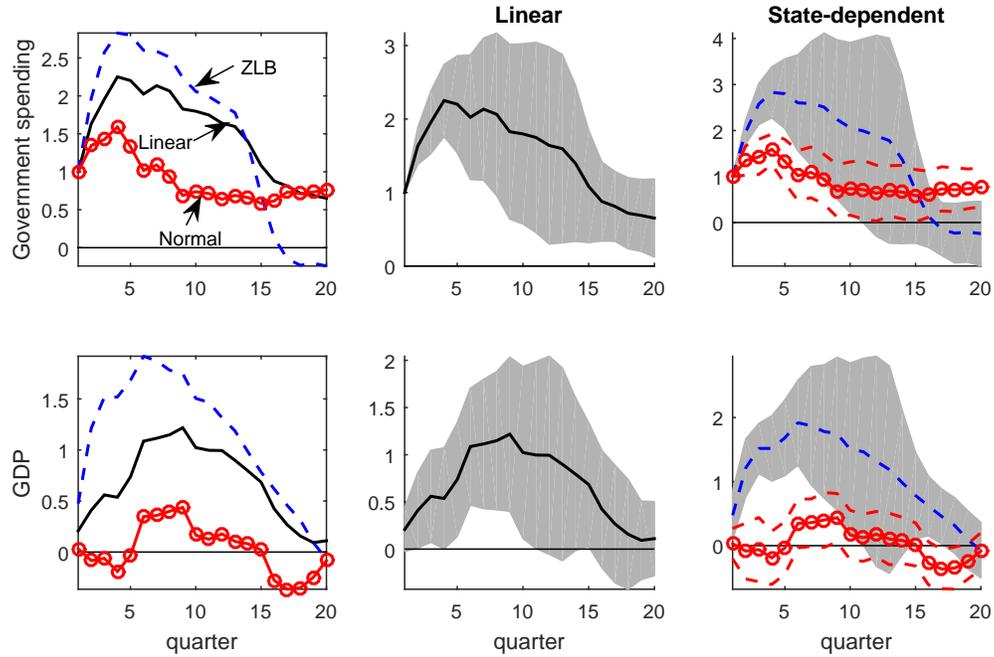
Note: These are responses for taxes and deficits in the linear model. The shaded areas indicate 95% confidence bands.

**Figure 6. State-dependent Taxes and Deficit Responses to a news shock: Considering slack states**



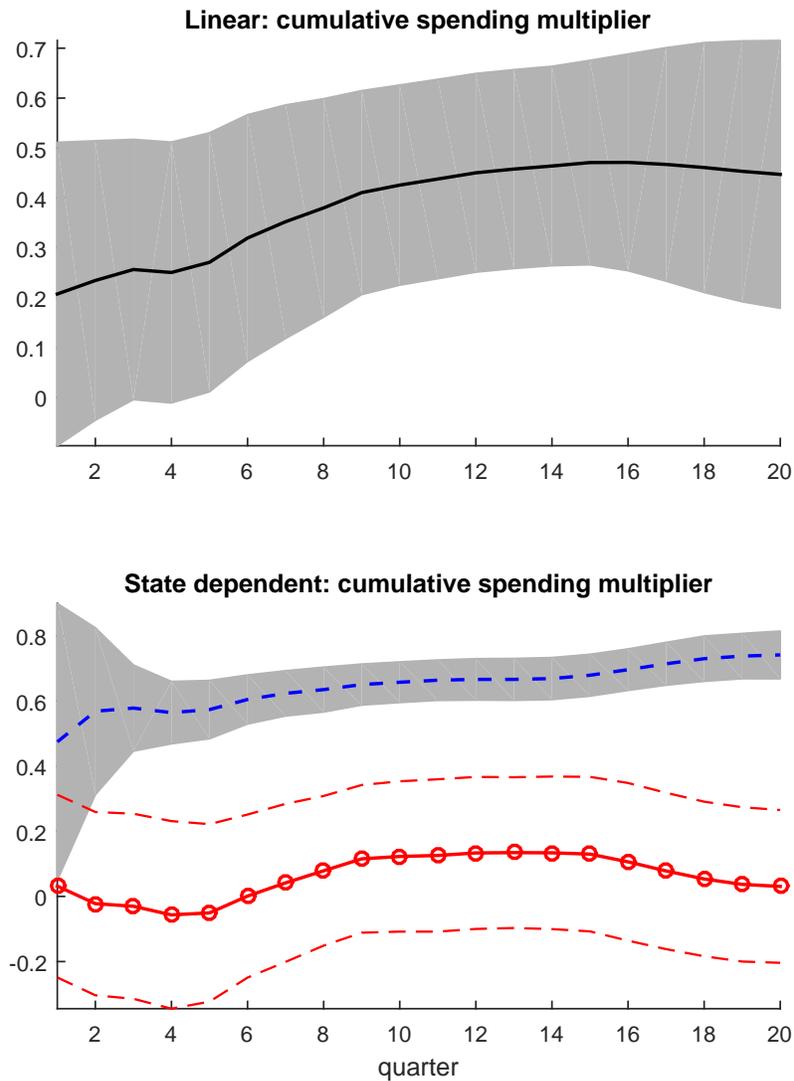
Note: These are state-dependent responses for taxes and deficits, where the black solid lines are responses in the high unemployment state and the lines with red circles are responses in the low unemployment state. 95% confidence intervals are also shown.

**Figure 7. Government spending and GDP responses to a Blanchard-Perotti shock: Considering zero lower bound**



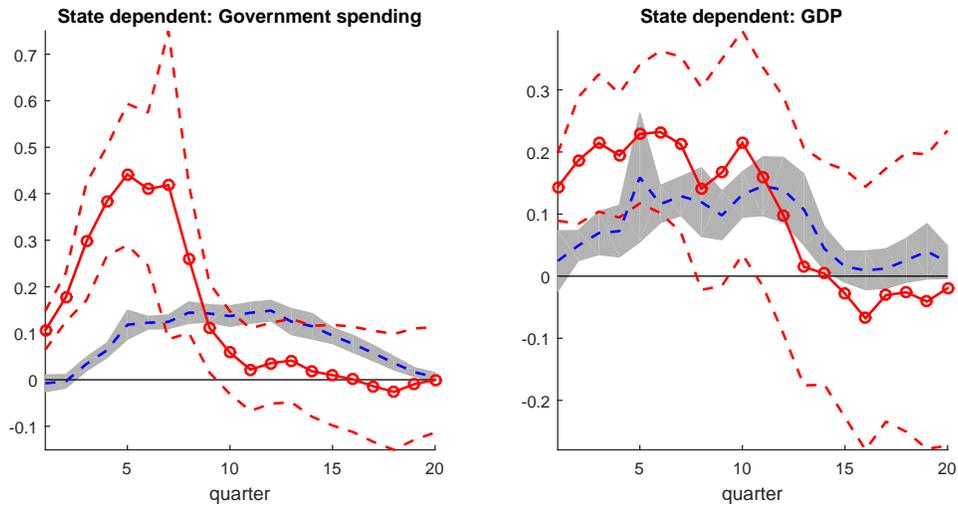
Note: Response of government spending and GDP to a 1% of government spending shock. The top row shows the response of government spending and the second row shows the response of GDP. The first column shows the responses in the linear and state-dependent model. The second column shows the responses in the linear model. The last column shows the state-dependent responses where the blue dashed lines are responses in the near zero-lower bound state and the lines with red circles are responses in the normal state. 95% confidence intervals are shown in second and third columns.

**Figure 8. Cumulative multipliers with Blanchard-Perotti shock: Considering zero lower bound**



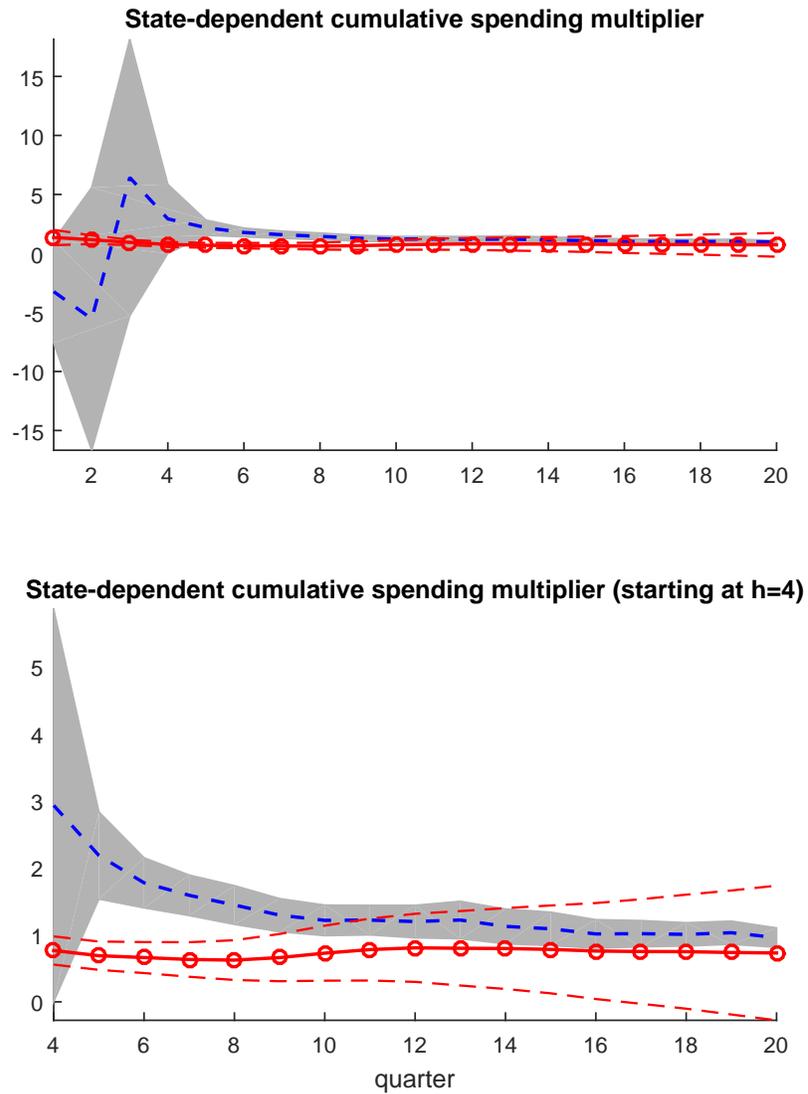
Note: Cumulative spending multipliers across different horizons for a Blanchard-Perotti shock. The top panel shows the cumulative multipliers in the linear model. The bottom panel shows the state-dependent multipliers where the blue dashed lines are multipliers in the near zero-lower bound state and the lines with red circles are multipliers in the normal state. 95% confidence intervals are shown in all cases.

**Figure 9. Government spending and GDP responses to a news shock: Considering zero lower bound and excluding World War II**



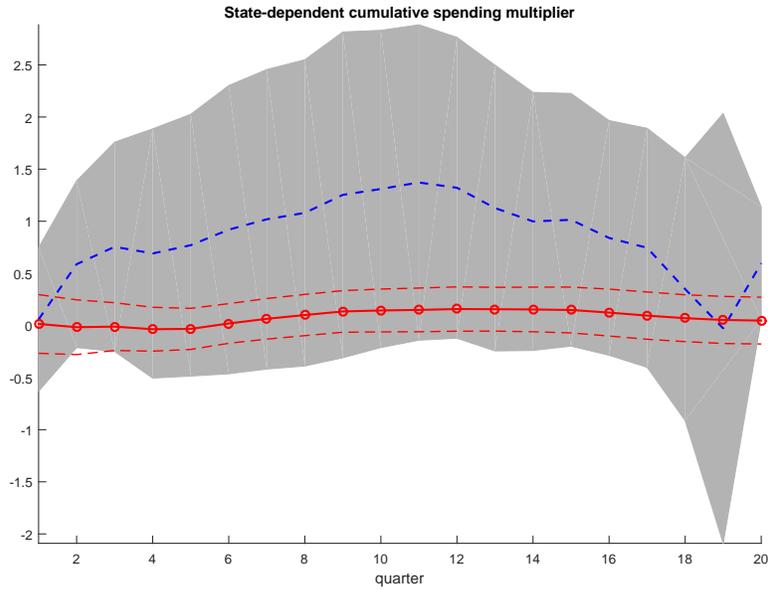
Note: Response of government spending and GDP to a news shock equal to 1% of GDP. The figure shows the state-dependent responses where the blue dashed lines are responses in the near zero-lower bound state and the lines with red circles are responses in the normal state. 95% confidence intervals are shown in all cases.

**Figure 10. Cumulative multipliers with news shock: Considering zero lower bound and excluding WWII**



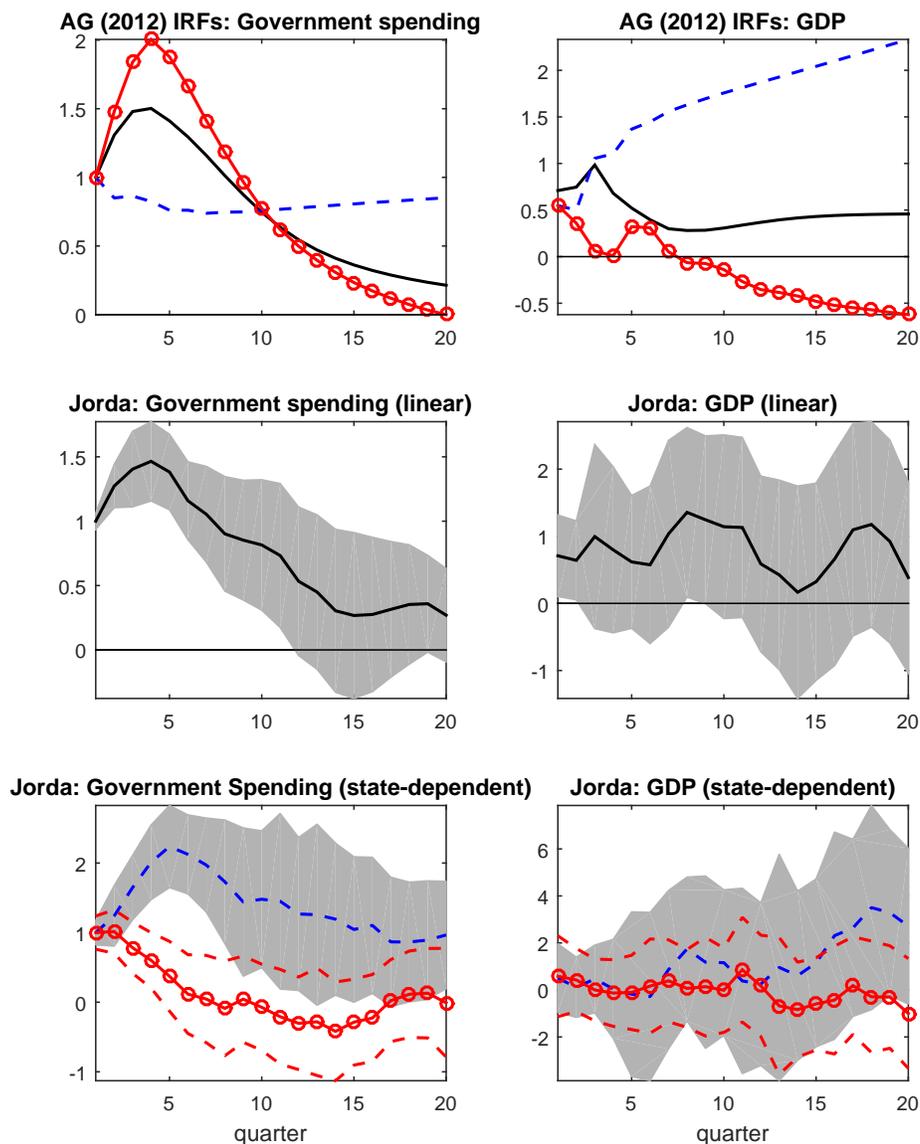
Note: Cumulative spending multipliers across different horizons for a news shock. The top panel shows the state-dependent multipliers where the blue dashed lines are multipliers in the near zero-lower bound state and the lines with red circles are multipliers in the normal state. The bottom panel shows the same figure but for horizon  $h \geq 4$ . 95% confidence intervals are shown in all cases.

**Figure 11. Cumulative multipliers with Blanchard-Perotti shock: Considering zero lower bound and excluding WWII**



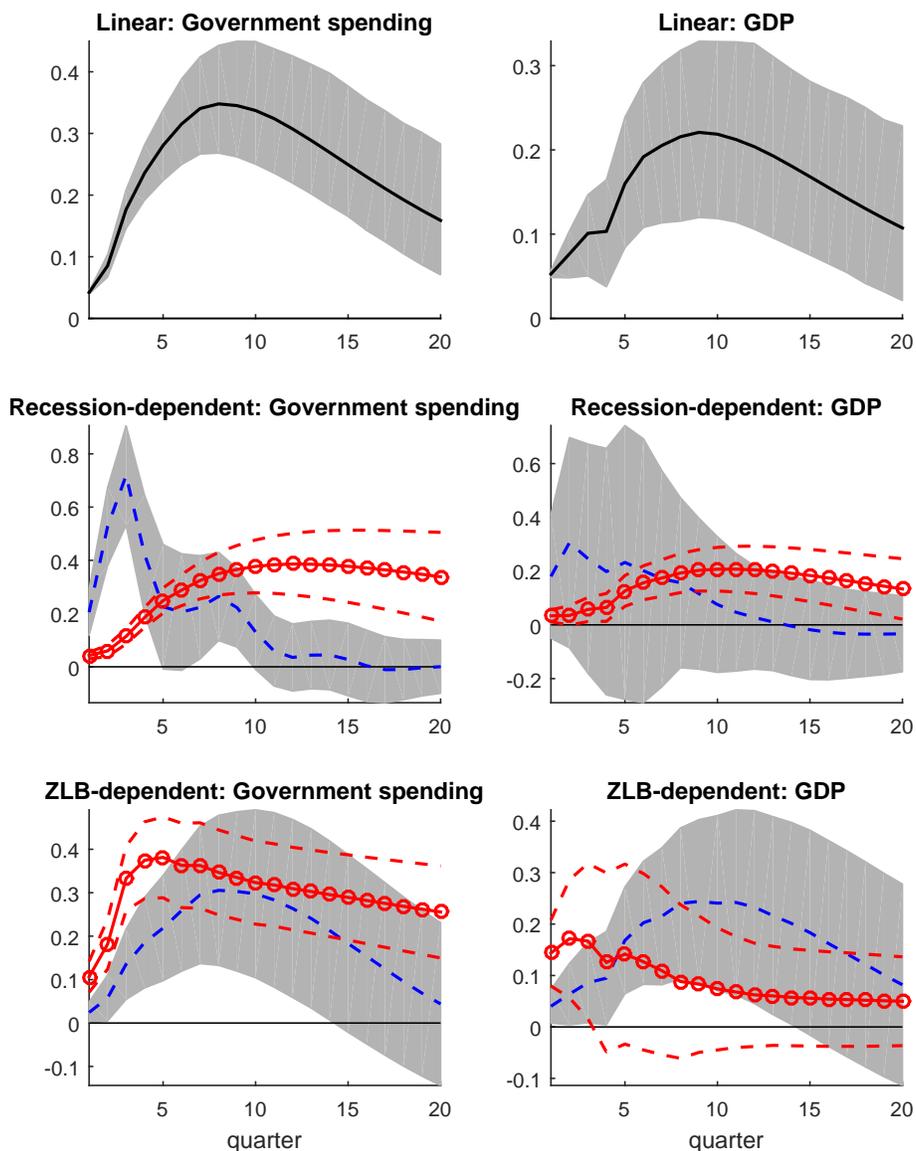
Note: Cumulative spending multipliers across different horizons for a Blanchard-Perotti shock. The figure shows the state-dependent multipliers where the blue dashed lines are multipliers in the near zero-lower bound state and the lines with red circles are multipliers in the normal state. 95% confidence intervals are shown in all cases.

**Figure 12. Estimating Auerbach and Gorodnichenko (2012) State Dependent Responses with the Jordà method**



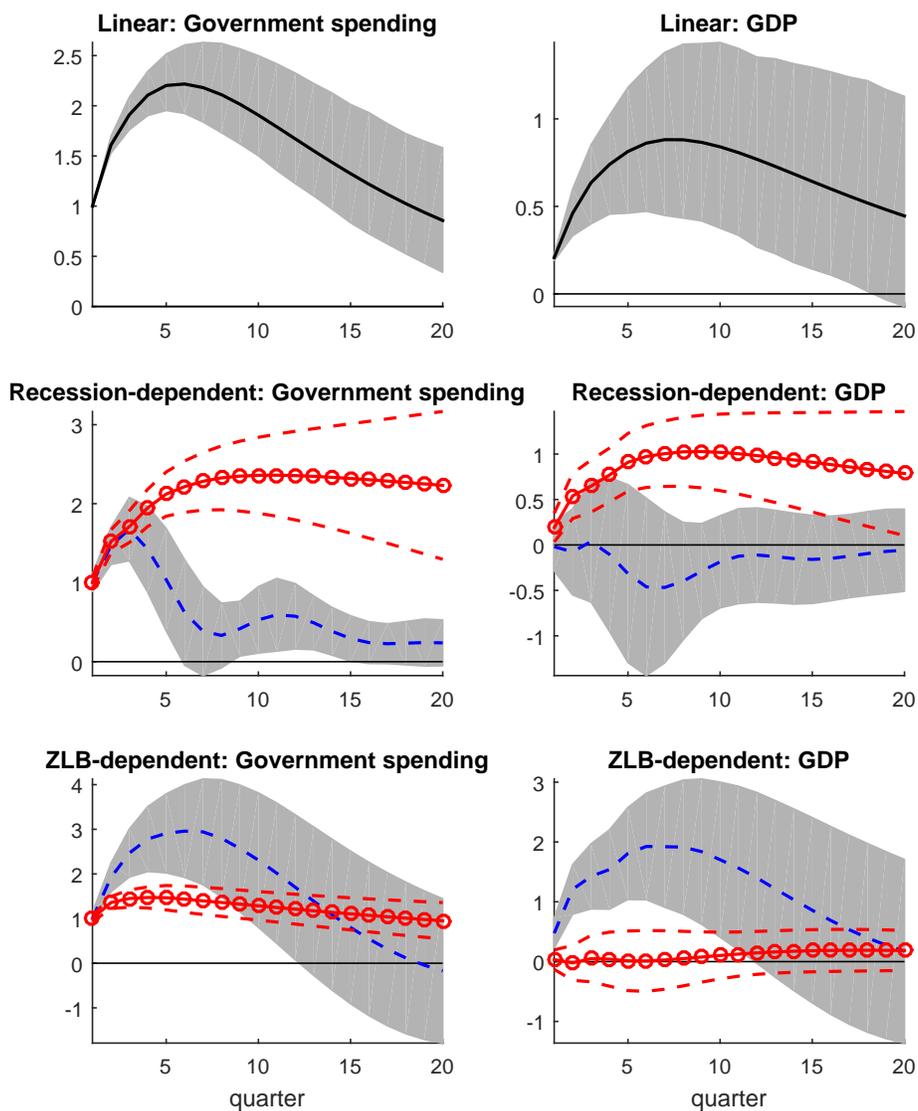
Note: The top panel replicates the responses for government spending and GDP from Figure 2 of AG(2012). These show the responses in the linear (black solid), recession (blue dashed) and expansion (red circles). The bottom two rows show the response of government spending and GDP to a government spending shock equal to 1% of GDP, with the same data, identification scheme and threshold definition as Auerbach and Gorodnichenko (2012), using the Jordà method. The second row shows the responses in the linear model. The last row shows the state-dependent responses in recession (blue dashed) and expansions (red circles). 95% confidence intervals are shown in all cases.

**Figure 13. Government spending and GDP responses to a news shock: Considering ZLB and Recessions in a Threshold-VAR**



Note: Response of government spending and GDP to a news shock equal to 1% of GDP for our full sample. The top panel shows the responses in the linear model. The middle panel shows the state-dependent responses where the blue dashed lines are responses in recession and the lines with red circles are responses in expansions. The bottom panel shows the state-dependent responses where the blue dashed lines are responses in near zero-lower bound state and the lines with red circles are responses in normal times. 95% confidence intervals are shown in all cases.

**Figure 14. Government spending and GDP responses to a Blanchard-Perotti shock: Considering ZLB and Recessions in a Threshold-VAR**



Note: Response of government spending and GDP to a government spending shock equal to 1% of GDP for our full sample. The top panel shows the responses in the linear model. The middle panel shows the state-dependent responses where the blue dashed lines are responses in recession and the lines with red circles are responses in expansions. The bottom panel shows the state-dependent responses where the blue dashed lines are responses in near zero-lower bound state and the lines with red circles are responses in normal times. 95% confidence intervals are shown in all cases.

## References

- Alloza, Mario**, 2014. “Is Fiscal Policy More Effective in Uncertain Times or During Recessions?” Working paper, University College London.
- Auerbach, Alan and Yuriy Gorodnichenko**, 2012. “Measuring the Output Responses to Fiscal Policy.” *American Economic Journal: Economic Policy* 4(2): 1–27.
- Barro, Robert J. and Charles J. Redlick**, 2011. “Macroeconomic Effects from Government Purchases and Taxes.” *Quarterly Journal of Economics* 126(1): 51–102.
- Baxter, Marianne and Robert G. King**, 1993. “Fiscal Policy in General Equilibrium.” *American Economic Review* 83(3): 315–334.
- Blanchard, Olivier and Roberto Perotti**, 2002. “An Empirical Characterization of the Dynamic Effects of Changes in Government Spending and Taxes on Output.” *Quarterly Journal of Economics* 117(4): 1329–1368.
- Caggiano, Giovanni, Efrem Castelnuovo, Valentina Colombo, and Gabriela Nodari**, 2015. “Estimating Fiscal Multipliers: News From A Non-linear World.” *Economic Journal* 125(584): 746–776.
- DeLong, J. Bradford and Lawrence H. Summers**, 2012. “Fiscal Policy in a Depressed Economy.” *Brookings Papers on Economic Activity* 44(1): 233–297.
- Drautzburg, Thorsten and Harald Uhlig**, 2015. “Fiscal Stimulus and Distortionary Taxation.” *Review of Economic Dynamics* 18(4): 894–920.
- Fazzari, Steven M., James Morley, and Irina Panovska**, 2015. “State-dependent effects of fiscal policy.” *Studies in Nonlinear Dynamics & Econometrics* 19(3): 285–315.
- Gali, Jordi, J. David Lopez-Salido, and Javier Valles**, 2007. “Understanding the Effects of Government Spending on Consumption.” *Journal of the European Economic Association* 5(1): 227–270.
- Hall, Robert E.**, 2009. “By How Much Does GDP Rise If the Government Buys More Output?” *Brookings Papers on Economic Activity* Fall: 183–236.
- Koop, Gary, Hashem M. Pesaran, and Simon M. Potter**, 1996. “Impulse Response Analysis in Nonlinear Multivariate Models.” *Journal of Econometrics* 74(1): 119–147.
- Ramey, Valerie A.**, 2013. “Government Spending and Private Activity.” In *Fiscal Policy After the Financial Crisis*, edited by Alberto Alesian and Francesco Giavazzi, pp. 19–55. University of Chicago Press.
- Romer, Christina D. and David H. Romer**, 2010. “The Macroeconomic Effects of Tax Changes: Estimates Based on a New Measure of Fiscal Shocks.” *American Economic Review* 100(3): 763–801.