Government Spending Multipliers in (Un)certain Times

Jan Philipp Fritsche†, Mathias Klein‡ and Malte Rieth§

This version: June 14, 2019

Abstract

We estimate the dynamic effects of government spending shocks, using probabilistic instruments that exploit time-varying volatility in US data. We find that the average government spending multiplier is significantly and persistently above one, driven by a crowding-in of private consumption and non-residential investment. We rationalize the results empirically through a contemporaneously countercyclical response of government spending and an efficient weighting of observations inversely to their error variance. We then show that the multiplier is significantly smaller when volatility is higher, consistent with theories predicting reduced effectiveness of surprise fiscal interventions in uncertain times.

Keywords: Fiscal policy, government spending multiplier, uncertainty, structural vector autoregressions, heteroskedasticity.

JEL classifications: C32, E62, H50.

* We thank Carlo Favero, Ethan Ilzetzki, Christophe Kamps, Helmut Lütkepohl, and participants of the CEF Milan and several seminars for helpful comments and suggestions.
† Humboldt Universität Berlin, e-mail: jfritsche@diw.de
‡ DIW Berlin; e-mail: mklein@diw.de
§ Corresponding author, DIW Berlin, Mohrenstraße 58, 10117 Berlin, Germany; e-mail: mrieth@diw.de
1 Introduction

There is renewed interest among researchers and policy makers in the effects of fiscal policy on macroeconomic activity. This topic is increasingly important because central banks have lost their ability to stimulate the economy through significant reductions in the policy interest rate. Despite the importance of this research question, there is a wide range of different estimates for the size of the government spending multiplier; that is, by how much does aggregate output increase following an exogenous increase in government spending. In addition, there is a discussion whether the effectiveness of fiscal policy interventions depends on the state of the economy (Auerbach and Gorodnichenko, 2012; Ramey and Zubairy, 2018).

In this article, we use an agnostic identification approach to estimate the dynamic effects of government spending shocks in the United States. For identification we exploit time-variation in the volatility of US data, which we model through a Markov switching process following Herwartz and Lütkepohl (2014). Several studies show that changes in volatility are a main feature of macroeconomic time series (Stock and Watson, 2002, Justiniano and Primiceri, 2008, Carriero et al., 2016, Diebold et al., 2017). We use these changes in volatility as a ‘probabilistic instrument’ (Rigobon, 2003) for government spending shocks to study how fiscal policy influences the macroeconomy.

Compared to other identification schemes frequently used in the fiscal policy literature, identification through heteroskedasticity offers three main advantages. First, unlike the recursive identification scheme (Blanchard and Perotti, 2002), it does not impose a timing restriction and allows for a contemporaneous response of government spending to output. Second, contrary to the narrative identification scheme (Barro and Redlick, 2011; Ramey, 2011), it does not require the availability of a valid external instrument for post-WWII data. Third, different to identification in linear models, it generalizes the estimation by weighting observations inversely to their sampling uncertainty, thereby increasing efficiency.

Furthermore, because the model is nonlinear, it allows for studying the state-dependent effects of government spending shocks across different volatility regimes. Several studies show that there exists a close connection between volatility and uncertainty, because in times of high volatility, expectation formation for households and firms is more difficult. In particular, a higher level of uncertainty increases the real option value of waiting before making investment decisions (Bernanke, 1983; McDonald and Siegel, 1986). Moreover, in a situation of increased uncertainty, firms might postpone hiring and capital decisions (Bloom, 2009; Bloom et al., 2018). Additionally, in uncertain times there is a stronger precautionary savings motive by consumers (Fernandez-Villaverde et al., 2011, Basu and Bundick, 2017). By allowing the effects of government spending shocks to depend on the
volatility regime, our identification approach enables us to test whether these theoretical predictions are supported by the data.

We find that exogenous changes in government spending are, on average, an effective tool to stimulate the economy. A positive government spending shock leads to a significant increase in output, private consumption, and investment. The crowding-in of private consumption is driven by a persistent and significant rise in both non-durable and durable consumption. The increase in investment is due to higher non-residential investment expenditures. Moreover, the unexpected fiscal stimulus leads to significantly higher inflation, short-term interest rates, employment, and hours worked.

We estimate an average government spending multiplier of 1.9 upon impact and a cumulative multiplier of 1.1 five years after the shock. This estimate is in line with the recent evidence (Caldara and Kamps, 2017; Chodorow-Reich, 2019), but larger than estimates typically found when applying timing-restrictions or exogenous changes in defense spending. We show that there are three main reasons for the differences across identification schemes. First, not restricting the contemporaneous response of government spending to output to zero implies that we can estimate this elasticity. We document a significantly negative response. As the size of the government spending multiplier is negatively related to this elasticity, a countercyclical spending response implies a larger multiplier. Second, military spending shocks produce smaller total government spending multipliers than general government spending shocks. Third, accounting for heteroscedasticity generalizes the estimation. It gives more weight to observations with low error variance compared to observations with high error variance. This increases efficiency.

Regarding state-dependency, we find strong empirical evidence for the hypothesis that the impact of government spending shocks varies across volatility regimes. The multiplier is significantly larger in the low volatility state compared to the high volatility state with a point estimate being twice a large. This result is consistent with the theoretical predictions and it suggests that the level of uncertainty in the economy significantly influences the effectiveness of government spending policy. It adds to evidence highlighting that more disagreement amongst US professional forecasters about future government spending reduces the impact of recursively identified fiscal spending innovations (Ricco et al., 2016) and that higher macroeconomic uncertainty lowers the effectiveness of narratively identified tax shocks (Bertolotti and Marcellino, 2017). Moreover, it rationalizes our finding of an average multiplier above one as the generalized model attaches larger weight to the more precisely estimated larger multiplier of the low volatility regime.

The rest of the paper is structured as follows. Section 2 describes the methodology and data. Section 3 presents our baseline results and shows robustness of the estimated average multiplier. In Section 4, we discuss our findings in the light of the literature. In Section 5, we allow for state-dependent effects of government spending shocks and
investigate how uncertainty regimes affect the fiscal transmission mechanism. Finally, Section 6 concludes.

2 Empirical methodology and data

In this section, we first present the empirical model and the identification strategy. We then describe the data and the estimation procedure.

2.1 The MS-SVAR model

The general $M$ state, $p$ lag reduced form Markov switching in heteroskedasticity structural vector autoregressive (MS-SVAR) model with $n$ variables is

$$y_t = c + A_1y_{t-1} + \cdots + A_py_{t-p} + u_t,$$  

(1)

where, in our case, $y_t = [g_t, x_t, r_t]'$ is the vector of endogenous variables with $g_t$ government spending, $x_t$ output, and $r_t$ the three-month Treasury Bill rate. Further, $c$ is a vector of constants, $A_i$ are parameter matrices with $i = 1, \ldots, p$, and $u_t$ is a vector of zero-mean reduced form errors. In standard SVAR models, a linear transformation is used to obtain the structural shocks, $\epsilon_t$, as $\epsilon_t = B^{-1}u_t$ or $B\epsilon_t = u_t$. Usually, $\epsilon_t \sim (0, I_n)$ and the reduced form covariance matrix is decomposed as $E[u_tu_t'] = \Sigma_u = BB'$. 

In our analysis, we follow Herwartz and Lütkepohl (2014) and assume that $E[u_tu_t'] = \Sigma_u(S_t)$. $S_t$ is a first order discrete valued Markov process that can take on $M$ different values, $S_t = 1, \ldots, M$, with transition probabilities given by $p_{kl} = P(S_t = l | S_{t-1} = k), k, l = 1, \ldots, M$. For estimation with maximum likelihood (ML), we assume that $u_t$ is normally and independently distributed conditional on a given state:

$$u_t|S_t \sim \text{NID}(0, \Sigma_u(S_t)).$$  

(2)

The normality assumption is not critical for the empirical analysis. If conditional normality is not fulfilled, then the estimation will simply be pseudo ML.

While the model is linear in a given state, it is nonlinear as a whole. Equations (1) and (2) imply that only the reduced form covariance matrix switches between states, as we are interested in the heteroskedasticity features of the data for identification purposes. Thus, we impose more regularity on the model than in the MS-SVAR model used by Sims and Zha (2006). Furthermore, although there is a finite number of states in practice, the model captures smooth transitions between them as the actual volatility is described as a mixture of states, each weighted with a certain probability that may be less than one.

We exploit these changes in the covariance matrix for the identification of structural shocks that are consistent with the statistical properties of the data. We consider two
states of the world – a high-volatility state and a low-volatility state. Then, we can decompose the reduced form covariance matrices as:

$$\Sigma_u(1) = BB' \quad \text{and} \quad \Sigma_u(2) = B\Lambda_2 B'$$

with $\Lambda_2 = \text{diag}(\lambda_{21}, \ldots, \lambda_{2n})$ containing only positive elements on the main diagonal and otherwise zeros. Lanne et al. (2010) show that if these elements are all distinct, the decomposition in (3) is unique apart from changes in the signs of the shocks and permutations of the $\lambda_{2n}$’s and corresponding orderings of the columns of $B$. In summary, if we order the $\lambda_{2n}$, assume that the structural shocks are orthogonal, have the same impact effects across states, and are normalized to have unit variance in the first state, we can uniquely identify the structural shocks through the linear transformation $\epsilon_t = B^{-1}u_t$.

The crucial assumption for point-identification of the full $B$-matrix is that the $\lambda_{2K}$’s are all different. As they can be interpreted as variance shifts of the structural shocks relative to the benchmark state (in our case the low-volatility state), having distinct $\lambda_{2K}$s means that the volatility shifts are not the same for all shocks. This assumption can be checked after estimation, which is an advantage over the conventional zero restriction or instruments used to identify government spending shocks, as the implied assumptions of the latter two approaches are typically just-identifying such that the data have no chance to speak against them. Furthermore, if the model is identified through heteroskedasticity, any such zero restrictions on $B$ become over-identifying and can be tested.

Figure 1 illustrates how identification through heteroskedasticity works. Both panels present observations describing the relationship between government spending and output under the assumption that government spending negatively depends on output (government spending is countercyclical) and output increases with government spending (government spending multiplier is positive). It is impossible to identify the output response to variations in government spending from the left panel alone, because every line fitted to the system will match the data equally badly. However, once the variance of the spending shock increases, as shown in the right panel, shifts of the spending equation increase relative to shifts of the output equation. The circle turns into an ellipse that is centered around the output equation such that the latter can be estimated from the data. The important condition is that the changes in the variances of the structural shocks are not proportional, as this would simply imply a widening of the circle. More generally, if there are distinct relative variance shifts for the other structural shocks as well, the full impact matrix can be point-identified.

In sum, our identification strategy will exploit that in time periods with high volatility in government spending, we learn more about the response of output to spending as the variance of spending shocks in these time periods is higher and spending shocks are more likely to occur. Technically, the covariance between spending and output increases and
this provides additional information in form of second moments that can be used for identification. Relying on this statistical identification approach is also convincing from an economic point of view. Several studies document that changes in volatility are a main feature of macroeconomic data (Stock and Watson, 2002, Justiniano and Primiceri, 2008, Carriero et al., 2016, Diebold et al., 2017). One prominent example is the change associated with the Great Moderation that started around the 1980s. Macroeconomic aggregates are more volatile before and fluctuate less during that period. Our approach makes use of such data features to identify structural shocks. This is less restrictive than using timing, sign, or external instrument restrictions. It does not restrict the size or the sign of the impact effects, nor does it require the availability of a valid instrument for post-WWII data.

2.2 Estimation, data, and model selection

We estimate the parameters in (1) by means of the expectation maximization (EM) algorithm developed for structural models by Herwartz and Lütkepohl (2014). Crucial for the analysis is to incorporate the regime-switching nature of the covariance matrix described in (2), given the restrictions in (3). We use the following concentrated out log likelihood function in the maximization step of the EM algorithm, which weighs observations according to their sampling uncertainty:

$$L(B, \Lambda_m) = \frac{1}{2} \sum_{m=1}^{M} \left[ \hat{T}_m \log(\det(\Sigma_u(m))) + \text{tr} \left( (\Sigma_u(m))^{-1} \sum_{t=1}^{T} \hat{\xi}_{mt|T} \tilde{\beta}_t \tilde{\beta}_t' \right) \right],$$

where $\xi_{mt|T}, m = 1, \ldots, M, t = 1, \ldots, T$ are the model smoothed probabilities, $T_m = \sum_{t=1}^{T} \xi_{mt|T}$, and the hat denotes estimated parameters from the previous iteration.
Once the EM algorithm converges, we obtain standard errors of the point estimates of the parameters through the inverse of the negative Hessian matrix evaluated at the optimum. We use them as a statistic to determine whether the estimated variances change significantly and by differing amounts across states. This is a requirement for identification. For the dynamic analysis, we compute bootstrapped impulse responses, following Herwartz and Lütkepohl (2014). Given the heteroskedasticity, classical residual bootstrapping may be problematic for generating reliable confidence intervals. Resampling needs to preserve the second order characteristics of the data. Therefore, we use a fixed design wild bootstrap with $u_t^* = \varphi_t \hat{u}_t$, where $\varphi_t$ is a random variable independent of $y_t$ following a Rademacher distribution. $\varphi_t$ is either 1 or –1 with probability 0.5.

Our baseline model consists of three endogenous variables, namely government spending (the sum of government consumption and government investment), GDP, and the three-month Treasury Bill rate. Several studies show that the conduct of monetary policy influences the macroeconomic effects of fiscal policy (Canova and Pappa, 2011; Davig and Leeper, 2011). Specifically, the extent to which monetary policy reacts to the consequences of fiscal spending shocks is an important determinant of their size. Therefore, in any study of the effects of government spending shocks, it is important to explicitly control for the stance of monetary policy which we do by including a short term interest rate in the vector of endogenous variables.¹ All series are from Ramey and Zubairy (2018). Nominal government spending and output are both divided by potential output such that they are measured in dollar equivalents. The baseline sample covers the period 1954Q1-2015Q4. The starting date avoids the years from 1945 to the Korean war, commonly considered to be special within post-WWII data from a fiscal point of view (Monacelli et al., 2010). We use five lags to account for potential seasonal patterns in the quarterly variables. This also increases the confidence in results for longer impulse response horizons.²

Table 1: Model selection. Notes: $L_T$ denotes the likelihood function evaluated at the optimum, $AIC = -2\log(L_T) + 2f$, and $SC = -2\log(L_T) + \log(T)f$, with $f$ the number of free parameters.

<table>
<thead>
<tr>
<th>Model</th>
<th>$\log(L_T)$</th>
<th>SC</th>
<th>AIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reduced form linear VAR(5)</td>
<td>1791.8</td>
<td>-3335.5</td>
<td>-3493.6</td>
</tr>
<tr>
<td>Reduced form MS(2)-VAR(5)</td>
<td>1899.8</td>
<td>-3459.1</td>
<td>-3675.7</td>
</tr>
</tbody>
</table>

Table 1 shows some specification statistics for the two-state Markov switching model. It is clearly preferred over a standard linear VAR according to the log-likelihoods and both types of information criteria. The latter is shown to work well to judge the performance

¹ In a latter section, we show that our main qualitative results are robust when controlling for fiscal foresight and when excluding the period during which the interest rate has reached its lower bound.

² All our qualitative results are robust when adjusting the lag length.
of MS models (Psaradakis and Spagnolo, 2006), whereas standard tests are problematic for this purpose as some parameters might not be identified under the null hypothesis of a smaller number of states than under the alternative (Hansen, 1992). The evidence against the linear model is strong. This suggests notable changes in volatility in the sample and that the identification of the structural shocks can be achieved by relying on the heteroskedasticity property of the data.

3 Government spending shocks and multipliers

3.1 Estimated volatility regimes and identification

We first analyze the estimated state-dependent reduced form covariances to see whether the model captures and separates the changes in volatility apparent in the data, according to Table 1, which are crucial for our identification strategy. This information also helps us interpret our endogenously identified regimes. Table 2 presents the estimated state covariances of the MS(2)-SVAR(5) model. It shows strong increases in volatility in state 2. The variances of the reduced form errors in the equations for public spending, output, and the interest rate increase by factors of 3, 5, and 17, respectively. We read this as further evidence that the sample is characterized by strong changes in volatility.

Table 2: Estimated state covariances of MS(2)-SVAR(5) model with $y_t = [g_t, x_t, r_t]'$.

<table>
<thead>
<tr>
<th></th>
<th>State 1: $\Sigma_u(1)$</th>
<th>State 2: $\Sigma_u(2)$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.024 0.033 0.001</td>
<td>0.079 0.073−0.016</td>
</tr>
<tr>
<td></td>
<td>0.024 0.228 0.028 0.062</td>
<td>0.073 1.164 0.280 1.055</td>
</tr>
</tbody>
</table>

The table also shows that the covariances increase substantially (in absolute value) in state 2; by about similar multiples as the variances relative to state 1. These changes in the covariances illustrate the idea behind identification through heteroskedasticity. In a period of high government spending volatility, we learn more about the relation between government spending and output as the covariance between both variables temporarily increases. Government spending shocks are then more likely to occur and can be used as a ‘probabilistic instrument’ (Rigobon, 2003) to trace out the response of output.

To achieve identification from a statistical point of view, we need significant and differential changes in the volatility of the structural shocks. Table 3 shows the estimated structural variances in state 2. As the ordering of the $\lambda_{2n}$’s is arbitrary, we simply order them from smallest to largest. All estimates are significantly larger than one, implying that all structural shock variances increase when switching from state 1 to state 2. Thus, we label state 2 the high volatility state. Identification requires that the variance shifts
are all distinct from each other. This is the case according to their standard errors, which do not overlap, and suggests that the decomposition in (3) is unique.

Table 3: Estimates and standard errors of relative variances of the MS(2)-SVAR(5) model. *Note*: The standard errors are obtained from the inverse of the negative Hessian evaluated at the optimum.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>Standard error</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\lambda_{21}$</td>
<td>3.25</td>
<td>0.89</td>
</tr>
<tr>
<td>$\lambda_{22}$</td>
<td>6.05</td>
<td>1.38</td>
</tr>
<tr>
<td>$\lambda_{23}$</td>
<td>17.28</td>
<td>3.75</td>
</tr>
</tbody>
</table>

Figure 2 presents the estimated smoothed state probabilities to develop a notion about the economic drivers of the regimes. The upper part shows the smoothed probabilities for the high volatility regime. The high volatility regime prevails in the first part of the sample, the low-volatility state dominates the second part. This general pattern might reflect the changes associated to the relatively tranquil times of the Great Moderation in the 1990s and 2000s with high growth and low inflation. In addition, many of the transitions to the high volatility state are associated with specific events in the economic history of the US. There are peaks around the OPEC oil prices shocks in the late 1960s and the beginning of the 1970s. Moreover, there is a spike around the Energy Crisis, the subsequent economic recession, as well as the chairmanship of the Federal Reserve of Paul Volcker at the end of the 1970s and the first half of the 1980s. In the second part of the sample, there are peaks around the burst of the Dot-com Bubble in 2001 and the Great Recession in 2009. Overall, this short narrative suggests that the endogenously determined regimes capture relevant developments in the US.

The estimated probabilities and shock variances can also be rationalized from a fiscal policy perspective. As shown by Davig and Leeper (2011), fiscal policy during the pre-Volcker period is characterized by a high degree of instability. In particular, they find that during that period, fiscal policy deviated substantially from a policy rule that induces government debt sustainability. This fiscal instability can affect private agents’ consumption and investment decision as it complicates the projection of future policy adjustments. Consequently, an unstable fiscal policy environment might increase economic uncertainty reflected in larger fluctuations of aggregate variables. Besides the specific conduct of fiscal policy, the smoothed probabilities spike around specific war episodes in the US history, such as the Lebanon Crisis, the Vietnam War, and the Afghanistan War. Because military spending is one main part of overall government spending, substantial increases in defense spending induced by military interventions lead to larger movements in government spending, which finally implies more volatility in public expenditures.

The bottom panel shows recessions, as defined by the NBER. It highlights that state 2 has some commonalities with recessions but measures something different. The correlation
Figure 2: Smoothed state probabilities. Notes: The figure shows the smoothed state probabilities \( \hat{\xi}_{mt|T} \) for the high volatility state \( m = 2 \) in the upper panel and recessions defined by the NBER in the lower panel.

between both series is roughly 0.43. Our latent Markov state is conceptually related to expectations and forecastability in a system of equations. When uncertainty is high, it is more difficult to make accurate predictions about the future, because the variables in the system are more volatile. The recession indicator mostly captures the current state of the economy and is agnostic about the future and expectations.

Having labeled the regimes, we now turn to the interpretation of the structural shocks. As mentioned above, our data-driven identification scheme is not based on a priori economic reasoning about the model-economy and a corresponding ordering of the columns of \( B \) or the respective \( \lambda \)'s. Thus, to attach an economic label to the statistically identified shocks, we inspect the forecast error variance decomposition and call the shock that explains the largest part of the variance in government spending a government spending shock. This idea is based on the identification scheme of Blanchard and Perotti (2002), which implies that government spending is contemporaneously exclusively driven by public spending shocks. We relax this assumption and allow for the possibility that other shocks affect government spending as well instantaneously, but we maintain the presumption that government spending shocks are the main driver at the quarterly frequency.

Table 4 shows the variance decompositions at various horizons and in both states. As before, shocks are ordered according to the size of their variances. Given our labeling scheme, there is little ambiguity in the interpretation of the structural shocks. We call the first shock the government spending shock, \( \epsilon^g_t \). It explains more than 95% of the variance of government spending at all horizons and in both states. This already indicates that our government spending shocks are similar to recursively identified ones, but not identical.
Table 4: Forecast error variance decompositions.

<table>
<thead>
<tr>
<th>Horizon</th>
<th>Variable</th>
<th>Shock in state 1</th>
<th>Shock in state 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>$\epsilon^g_t$</td>
<td>$\epsilon^{tfp}_t$</td>
</tr>
<tr>
<td>1 quarter</td>
<td>Spending</td>
<td>0.96 0.04 0.00</td>
<td>0.93 0.07 0.00</td>
</tr>
<tr>
<td></td>
<td>GDP</td>
<td>0.37 0.62 0.01</td>
<td>0.24 0.73 0.03</td>
</tr>
<tr>
<td></td>
<td>T-Bill rate</td>
<td>0.01 0.02 0.98</td>
<td>0.00 0.01 0.99</td>
</tr>
<tr>
<td>4 quarters</td>
<td>Spending</td>
<td>0.96 0.04 0.00</td>
<td>0.91 0.07 0.02</td>
</tr>
<tr>
<td></td>
<td>GDP</td>
<td>0.38 0.61 0.01</td>
<td>0.24 0.73 0.03</td>
</tr>
<tr>
<td></td>
<td>T-Bill rate</td>
<td>0.04 0.08 0.88</td>
<td>0.01 0.03 0.96</td>
</tr>
<tr>
<td>12 quarters</td>
<td>Spending</td>
<td>0.97 0.03 0.01</td>
<td>0.92 0.05 0.03</td>
</tr>
<tr>
<td></td>
<td>GDP</td>
<td>0.39 0.58 0.02</td>
<td>0.24 0.68 0.08</td>
</tr>
<tr>
<td></td>
<td>T-Bill rate</td>
<td>0.07 0.20 0.73</td>
<td>0.02 0.09 0.90</td>
</tr>
<tr>
<td>20 quarters</td>
<td>Spending</td>
<td>0.98 0.02 0.00</td>
<td>0.95 0.03 0.02</td>
</tr>
<tr>
<td></td>
<td>GDP</td>
<td>0.38 0.56 0.06</td>
<td>0.22 0.61 0.17</td>
</tr>
<tr>
<td></td>
<td>T-Bill rate</td>
<td>0.11 0.23 0.66</td>
<td>0.03 0.10 0.87</td>
</tr>
</tbody>
</table>

We confirm this impression in Figure 3, which compares the standardized government spending shocks to those of a linear model identified through a Cholesky decomposition with government spending ordered first. The correlation between both series is 0.85, suggesting that the shocks are closely related but different. These differences will turn out to be important quantitatively, as we show below. The shocks from the linear model show a clear pattern of heteroskedasticity, being generally more volatile during the first part of the sample and fluctuating less in the second part. This coincides with the estimated volatility states in Figure 2, which are mostly in the high volatility regime before the mid-1980s. Thus, in contrast to the MS model, the linear model does not appropriately account for the volatility changes of the underlying data.

Finally, we label the second shock a technology shock, $\epsilon^{tfp}_t$. It explains more than half of the variation in output in both states and at all horizons. We call the last shock a monetary policy shock, $\epsilon^mp_t$, as it explains nearly all of the unpredictable changes in the T-Bill rate in the first year, and at least 66% after five years.

### 3.2 Impulse responses for baseline model

Figure 4 shows the impulse responses to all three shocks. The first column presents the effects of a government spending shock on the endogenous variables in rows. Shaded areas indicate 90% confidence bands. A shock of one standard deviation leads to a significant contemporaneous increase in government spending by about 0.15%. The response is highly persistent, reaching a maximum after three years. The unexpected fiscal expansion has an immediate and substantial positive effect on output, which increases significantly by

---

3 The responses are calculated relative to state a (low volatility). Because the estimated coefficients do not vary across states, the responses in state b (high volatility) are just up to scaling.

4 Although highly persistent, the spending responses converges back to steady state at longer horizons.
0.3%. GDP rises further for the first three quarters, before gradually returning to the level where it would have been without the shock. The T-Bill rate increases significantly upon impact as well and remains statistically different from zero for about one year. The expansionary effect induced by the fiscal stimulus leads to an increase in interest rates by roughly 10 basis points.\(^5\)

The second column summarizes the impact of a technology shock. Output increases significantly by 0.4% in the same quarter the shock hits, and then in a hump-shaped manner. The peak response is at three quarters. Government spending falls upon impact. Although this effect is small and only borderline significant, the countercyclical behavior of government spending is of central importance for the size of the government spending multiplier (Caldara and Kamps, 2017). The government spending response is insignificant for the remaining horizon. Furthermore, the T-Bill rate increases significantly and persistently. The third column presents the responses to a monetary policy shock. After a short rise in output, an exogenous increase in the T-Bill leads to a persistent and significant decline in GDP. The government spending response is estimated to be insignificant.

A key concept in the fiscal policy literature is the government spending multiplier. It is also of great importance for the policy debate because it gauges the effectiveness of surprise fiscal stimuli. It measures by how many dollars aggregate output increases when government spending increases exogenously by one dollar. A multiplier above one implies that an unexpected increase in government spending crowds-in private demand, whereas a multiplier below unity implies a crowding-out. We focus on the dynamics of

\(^5\) As shown below, the government spending shock also leads to an increase in inflation, such that the interest rate response can be reconciled with a monetary authority following a standard Taylor-rule.
the cumulative government spending multiplier, which measures the cumulative change in output relative to the cumulative change in government spending from the time of the government expenditure shock to a given period of the forecast horizon. We compute it as the ratio between the cumulative output and the cumulative government spending response as both variables are already normalized by potential output.

Figure 5 shows the estimated cumulative multiplier together with 90% confidence bands. The multiplier is larger than one over the whole forecast horizon, and statistically
larger than unity for about ten quarters. It is close to two upon impact and it peaks two quarters after the shock. Subsequently, it decreases gradually and it is 1.1 after five years. Other studies find similar multiplier values (Ben Zeev and Pappa, 2017; Chodorow-Reich, 2019), but our estimate is larger than those typically found by applying timing-restrictions or using narrative accounts (Blanchard and Perotti, 2002; Ramey, 2011; Ramey and Zubairy, 2018). In the next section, we show empirically that there are three main reasons for this difference.

One of the reasons is based on the insights of Caldara and Kamps (2017). The authors show that the size of the multiplier depends negatively on the contemporaneous elasticity of government spending to output. The intuition for this relation can be summarized as follows. There exists a positive comovement between government spending and output in the data, which any identification approach decomposes into a fraction explained by government spending shocks and a fraction explained by the remaining shocks of the SVAR. The timing assumption of Blanchard and Perotti (2002) implies that government spending does not contemporaneously respond to any other shock. Therefore, all the positive comovement in the data must be explained by the government spending shock. This leads to a spending multiplier of about one. If the systematic response of government spending increases, the remaining shocks explain a larger part of the positive comovement and the spending multiplier decreases. In contrast, if the systematic response decreases, that is, turns negative, the multiplier increases.

This is what we find. Our estimate suggests a significant contemporaneous response of government spending to output of $-0.08$ and, thereby, a relatively large multiplier. Caldara and Kamps (2017) estimate an impact reaction of spending to output of $-0.13$ and obtain spending multipliers larger than one, as well. They explain the negative reaction by nominal government spending being not fully indexed to inflation, implying that real government spending falls in response to an increase in inflation induced by the surprised output expansion.

While Caldara and Kamps (2017) rely on external instruments to estimate the government spending to output elasticity and the corresponding multiplier, we use the heteroskedasticity in the data for estimating these relationships. We see our approach as complementing theirs. At the same time, ours gives full voice to the data, as it does not have to assume instrument exogeneity. Furthermore, it is not plagued by weak instruments problem for post-WWII data, it does not require the time-consuming construction of narrative instruments, and it can be readily applied to other countries.

3.3 Economy-wide effects

In this subsection, to understand which components of private demand drive our multiplier estimates, we analyze how government spending shocks affect consumption and
investment. Moreover, we discuss the effects on hours worked, employment, inflation, and
government deficit. For this, we augment our baseline model by one of the additional
variables at a time and summarize their impulse responses in Figure 6.

Figure 6: Impulse responses of other variables of interest to a government spending shock. Notes:
90% confidence bands constructed by a wild bootstrap.

Private consumption expenditures increase significantly. The maximum response is
reached two quarters after the shock occurred. This increase in consumption is driven
by a significant and quantitatively similar rise in both of its components. Households
purchase more non-durable and durable goods. The increase in private consumption
following an exogenous increase in public consumption is difficult to rationalize using
standard textbook models of the real business cycle or the New Keynesian paradigm.
In these models, a government spending increase that is financed through higher lump-
sum taxes in the future induces a negative wealth effect such that households increase
savings and reduce consumption expenditures. However, extensions of the baseline model,
such as limited asset market participation, countercyclical markups, or consumption-
leisure complementarity might overturn these effects such that households consume more
in response to a fiscal stimulus (Gali et al., 2007; Gnocchi et al., 2016; Ravn et al., 2012).

We use the Ramey and Zubairy (2018) dataset for the government deficit and the GDP Deflator.
Inflation is calculated as the quarterly percentage change of the GDP-Deflator. Total consumption
expenditures are calculated as the sum of the FRED-codes PCND, PCESV, PCDG, where, again, the
first two are used to calculate the non-durable consumption expenditures and the last one, PCDG, is
our measure of durable consumption expenditures. For investment we use the codes FPI and PNFI,
PRFI for (non-)residential investment, respectively. Data on total hours worked and employment in the
nonfarm business sector are taken from Valerie Rameys’ homepage. We normalize private consumption
and investment expenditures by potential output. Hours worked and employment are normalized by
total population and the government deficit is expressed relative to nominal GDP.
Our empirical evidence supports these theoretical modifications of the textbook model that limit the wealth effect due to an increase in government spending.\footnote{In general, the response of consumption to government spending shocks is the subject of a considerable debate with different results emerging from different identification schemes based on short-run restrictions and narrative identification schemes (Gali et al., 2007; Ramey, 2011).}

Private Investment rises significantly only in the first couple of quarters. After about one year, the response is insignificant and turns significantly negative – although only borderline – after about three years. The negative investment response at longer horizons is mainly due to a significant fall in residential investment, whereas non-residential investment is above trend for about three years and significantly so for six quarters. Although hard to reconcile with theory, the (short-lived) positive investment response is in line with the empirical evidence presented by Ben Zeev and Pappa (2017). Taken together, the large multiplier in the first year of the surprise fiscal expansion seems to be driven by a crowding-in of household consumption and firm non-residential investment.

Furthermore, the government spending shock leads to a significant increase in hours worked, which reaches a maximum after two quarters. Besides its effect on the intensive margin, the spending shock also affects the extensive labor margin. Total employment increases significantly and the response shows a similar pattern as the one of hours worked. These responses are perfectly in line with the predictions of textbook models in which households supply more labor to compensate for future tax increases due to the fiscal expansion. The spending shock also has a positive impact on inflation, which increases significantly and persistently. The inflationary effect rationalizes the persistent increase in the T-Bill rate in response to a government spending shock documented in Figure 4 as the endogenous response of monetary policy. Finally, the public deficit significantly increases, suggesting that the identified fiscal surprise expansions are mainly deficit-financed.

3.4 Robustness

We now show that our results are robust to various alterations of the model and the data.

**Augmented models.** Figure 7 contains the multiplier corresponding to each of the ten models underlying Figure 6. It shows that adding other variables to the model leaves the multiplier largely within the 90% confidence bands of the baseline estimate. Moreover, all multipliers are larger than one, suggesting a crowding-in of private demand for at least 12 quarters.

**Fiscal foresight.** As alternative sensitivity test, we account for fiscal foresight. This can arise when private agents do not just react to actual spending increases, but also to news about impending future spending plans. Then, the econometrician might not be able to recover the true unexpected spending shock because the information sets of agents and the econometrician are misaligned (Leeper et al., 2013). We address this issue by including
Figure 7: Multipliers of in augmented models. Notes: The figure shows the point estimates for the dynamic government spending multiplier of the augmented models (thin solid lines) and of the baseline model (thick line) and the 90% confidence bands for the baseline estimate.

a variable into the baseline SVAR that captures expectations about future policy actions. Specifically, we use real-time professional forecasts for government spending, which is a spliced series of government spending forecasts provided by the Greenbook (1966-2004) and the Survey of Professional Forecasters (1982-2016). We extend the series provided by Auerbach and Gorodnichenko (2012), which covers 1966 to 2008, to include the Great Recession and the following years. Figure 8 shows that the results (dashed line) are similar to the baseline estimates of the dynamic spending multiplier. In particular, the multiplier is larger than one. This analysis suggests that our baseline findings are quantitatively not strongly affected by fiscal foresight.

Zero lower bound. The baseline sample includes the Great Recession and the subsequent zero lower bound period. To investigate whether the results are robust to excluding these years, we estimate the baseline model ending the sample in 2007Q4. As the dotted line shows, the results are hardly affected when excluding the crisis and zero lower bound period. The estimated multiplier is within the confidence interval of the baseline estimate.

Tax policy. As the financing side of the fiscal stimulus might affect the size of the government spending multiplier, we control for tax policy in the estimation. In particular, the vector of observables now additionally includes real net tax revenues. The asterisked line presents the corresponding estimate and indicates that the main finding is not affected when controlling for tax policy.
4 Results in perspective

In this subsection, we compare our results to the literature and analyze the implications of our modeling choices for the size of the estimated multiplier. Our framework differs from Blanchard and Perotti (2002) and Ramey (2011) along two important dimensions. First, we use a different identification scheme, which is based on heteroskedasticity, whereas they use either a timing assumption or news on military spending as instrument. Second, our identification strategy implies that we estimate a nonlinear model that takes into account changes in the variances of the error terms, while they use linear models.

To quantify the importance of each dimension for explaining our results, we concentrate on the timing assumption first. It is usually justified by the fact that policy makers need time to decide on, approve, and implement changes in fiscal policy. To assess whether it is supported by the data, we set the contemporaneous response of government spending to technology shocks to zero, that is, we set $b_{12} = 0$ in restriction $R1$, and estimate this model against one with unrestricted impact matrix $B$. Table 5 displays the $p$-value of the corresponding likelihood-ratio (LR) test. The data do not reject $R1$. However, the test result needs to be treated with caution as the power of the test is unclear in our sample. Moreover, the evidence in Table 5 conflicts with Figure 4, which shows a borderline significant negative response of government spending to technology shocks. Taken together, we thus conclude that the data do not speak clearly in favor or against the timing assumption, in line with the evidence in Bouakez et al. (2014) and Lewis (2018).
Table 5: LR tests. Note: The table shows the null and alternative hypothesis for tests of $R_1$ and $R_2$ on the $B$-matrix and the associated degrees of freedom (df) and $p$-values.

<table>
<thead>
<tr>
<th>Null hypothesis</th>
<th>Alternative hypothesis</th>
<th>df</th>
<th>$p$-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$R_1 : b_{12} = 0$</td>
<td>$B$ unrestricted</td>
<td>1</td>
<td>0.382</td>
</tr>
<tr>
<td>$R_2 : b_{12} = b_{13} = b_{23} = 0$</td>
<td>$B$ unrestricted</td>
<td>3</td>
<td>0.573</td>
</tr>
<tr>
<td>$R_2 : b_{12} = b_{13} = b_{23} = 0$</td>
<td>$R_1 : b_{12} = 0$</td>
<td>2</td>
<td>0.539</td>
</tr>
</tbody>
</table>

In the next step, we impose two more restrictions that help us approaching a just-identified model in a linear setting; that is, without heteroskedasticity, against which we can compare the effect of modeling changes in volatility explicitly. In $R_2$, we impose additional zeros on the impact reaction of government spending and output to the monetary policy shock, that is, we set $b_{12} = b_{13} = b_{23} = 0$. This implies a Cholesky decomposition. The additional restrictions are motivated by a literature on the identification of monetary policy shocks (Christiano et al., 1999), which assumes that prices are rigid and that the real economy is affected by such shocks with a time lag. The data do not reject $R_2$, neither against the alternative of unrestricted $B$, nor against $R_1$ (see Table 5).

Table 6 shows the estimated cumulative output multiplier corresponding to the baseline model with unrestricted $B$ and to the model under $R_1$ and $R_2$, respectively, for selected years of the response horizon. In line with the argument of Caldara and Kamps (2017), the estimated multiplier decreases for all years under $R_1$. The decline is substantial. The estimate for the first year drops from 2.15 to 1.43. For year 5, the unrestricted model suggests a crowding-in, while MS-R1 implies a crowding-out of private demand with a multiplier below unity. In contrast, the two additional exclusion restrictions in MS-R2 have essentially no extra effect on the dynamic multiplier. Thus, together with the evidence from the LR-tests, we maintain $R_2$ and estimate a just-identified linear model. The estimates for the dynamic multiplier shown in the next row drop further when neglecting the heteroskedasticity. It declines to one in the first year and falls below unity for the remaining years. Summing up, the baseline unrestricted MS model yields a multiplier that is roughly twice as large as the one implied by a linear-recursive model. About two-thirds of that difference can be explained by differences in the identification scheme and the remaining fraction by alternative ways of dealing with changes in volatility.

Next, we compare the baseline estimate for the government spending multiplier to the output effects of military spending shocks, both in a heteroskedastic and in a linear model, to understand how the size of the multiplier depends on the public spending component considered and whether the finding of a larger multiplier in a time-varying volatility model generalizes to this type of shocks. A number of studies use a narrative approach (Ramey, 2011, Barro and Redlick, 2011 and Ramey and Zubairy, 2018). In this literature, wars and the associated expenses are assumed to be exogenous to current economic conditions. We follow this idea in our framework and augment the three-variable model with military
Table 6: Comparison of total government spending multipliers. Note: The table shows the cumulative total government spending multiplier for Markov switching (MS) and linear models for different model specifications and identification schemes. In the MS baseline model, the impact matrix $B$ is unrestricted. For $R1 : b_{12} = 0$ and for $R2 : b_{12} = b_{13} = b_{23} = 0$.

<table>
<thead>
<tr>
<th>Model</th>
<th>Response horizon in years</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td><strong>Government spending shock</strong></td>
<td></td>
</tr>
<tr>
<td>MS unrestricted</td>
<td>2.15</td>
</tr>
<tr>
<td>MS-R1</td>
<td>1.43</td>
</tr>
<tr>
<td>MS-R2</td>
<td>1.40</td>
</tr>
<tr>
<td>Linear-R2</td>
<td>1.01</td>
</tr>
<tr>
<td><strong>Military spending shock</strong></td>
<td></td>
</tr>
<tr>
<td>MS-R2</td>
<td>0.92</td>
</tr>
<tr>
<td>Linear-R2</td>
<td>0.78</td>
</tr>
</tbody>
</table>

spending, which we order first. In other words, we assume that military spending is contemporaneously exogenous with respect to the three baseline variables, for which we maintain the recursive identification. This assumption is also applied by, among others, Barro and Redlick (2011) and Hall (2009).

The bottom two rows of Table 6 show that the total government spending multiplier induced by military spending shocks is smaller than the total government spending multiplier generated by total government spending shocks, and below one throughout the full response horizon. As discussed in the literature, this might be due to the different nature of these shocks, which are typically not tailor-made to stimulate the economy, to the observation that large increases in military spending are often accompanied with other macroeconomic policies, like price controls, or to other confounding developments such as rising patriotism. Moreover, the comparison between both models for military spending shocks shows that the MS model implies a larger multiplier throughout the five years.

The second finding – together with the same observation for the multiplier produced by total government spending shocks – deserves further attention. It suggests that the estimated multiplier does not only depend on the identification scheme but also on the way the model deals with heteroskedasticity. Specifically, it seems that the generalized models, which account for time-varying volatility, imply larger estimates than the linear models, which ignore heteroskedasticity. In analogy to standard regression analysis, this outcome is intuitive as generalized models weigh observations differently than homoskedastic models. While the former give more weight, or likelihood (see Eq. 4), to observations with low error variance than to observations with high error variance, the latter treat them equally. This difference does not just affect inference but also point estimates.

The reasoning opens the door for two possible explanations for the larger multipliers implied by the heteroskedastic models. First, it could be that the multiplier is the same in both volatility regimes and that the larger estimates for the MS models are only due to
sampling error in the high volatility regime, which erroneously implies a lower multiplier for these observations; luckily, the estimate is down-weighted because of the associated higher uncertainty. If this is the case, it seems reasonable to use the estimates from the more efficient MS model. Lütkepohl and Schlaak (2018) show that heteroskedastic VAR models typically outperform linear models in terms of estimation precision of the structural parameters when the data feature changes in volatility.

Second, it may be that the multiplier is state-dependent. In particular, if the multiplier is larger in the low than in the high volatility regime, a Markov switching in heteroskedasticity model with state-independent impact matrix would yield a larger average multiplier than a linear model. It would give more weight to the larger multiplier in the low volatility state as this estimate is associated with lower sampling uncertainty. We investigate the second possibility next.

5 The state-dependent effects of government spending shocks

In this section, we relax the assumption of constant impact effects across volatility regimes. We allow for state-dependent effects to analyze whether the dynamic impact of government spending shocks depends on the level of aggregate uncertainty in the US economy. This is a central implication of seminal theoretical work, which shows that demand policy is less effective when volatility is high. In particular, Bernanke (1983) and McDonald and Siegel (1986) show that a higher level of uncertainty increases the real option value of waiting before making investment decisions. Moreover, in a situation of increased uncertainty, modeled through larger structural shock variances, firms might postpone hiring and investment decisions (Bloom, 2009; Bloom et al., 2018). Additionally, Basu and Bundick (2017) and Fernandez-Villaverde et al. (2011) point out that, in uncertain times, there is a stronger precautionary savings motive by consumers.

The Markov switching in heteroskedasticity model provides a natural way to test the prediction that fiscal stimulus is less effective during periods of high uncertainty as it allows for time-varying variances. Moreover, the framework provides advantages over models with exogenously determined regimes, such as threshold or smooth transition models, as it treats the potential transition variable(s) as latent. It is more agnostic about the state determination and reduces the risk of misspecification of the transition variable, function, or points. Relative to threshold models it is more flexible as it allows for mixtures of states.8

8 A related literature uses threshold or smooth transition models to analyze whether the impact of fiscal policy depends on the slack in the economy or on the level of the monetary policy rate (Auerbach and Gorodnichenko, 2012; Ramey and Zubairy, 2018).
To implement the dependency of the multiplier on the volatility regime, we generalize the decomposition in (3) as in Bacchiocchi and Fanelli (2015) to

\[ \Sigma_u(1) = BB' \quad \text{and} \quad \Sigma_u(2) = (B + Q)\Lambda_2(B + Q)', \]

(5)

where \( Q \) is a \( n \times n \) coefficient matrix that is added to the impact matrix \( B \). Now, the observable changes in the reduced form covariance matrix between regimes can be explained by changes in the impact effects, by changes in volatility of the structural shocks (as before), or by a combination of both. The decomposition in (3) is a special case of (5) with the restrictions \( Q = 0_{n \times n} \). As this special case was just-identifying \( B \), it follows that the decomposition (5) is not (locally) unique and we need to place some restrictions on the elements of \( B, Q \) and \( \Lambda_2 \). Specifically, the decomposition contains \( n(n + 1) \) symmetry restrictions but \( 2n^2 + n \) elements, such that we need \( n^2 \) further restrictions.

These are driven by our research question. As we are particularly interested in the state-dependent effects of government spending shocks, we keep the parameters mostly related to monetary policy constant and let the coefficients relevant for fiscal policy change across regimes. In detail, we set \( b_{13} = b_{23} = 0 \), based on the LR-tests in Table 5. In contrast, we leave \( b_{12} \) unconstrained. Although the LR-test does not reject the zero assumption for this parameter, the confidence bands of the impulse response of government spending do (see Figure 4). Moreover, Table 6 and Caldara and Kamps (2017) show that small differences in this parameter can have large effects on the size of the multiplier.

Regarding variances, we leave the volatility change of monetary policy shocks, \( \lambda_{23} \), unrestricted, given that we find the largest change in variance for this shock (see Table 3). For the other two shocks, we instead restrict the variances to be state-invariant and shift the changes in shock size to the corresponding diagonal elements \( q_{11} \) and \( q_{22} \) in \( Q \). We do this purely for computational reasons as this decomposition produces more stable results with the EM algorithm than a model where the main diagonal of \( Q \) is set to zero and the one of \( \Lambda_2 \) is left unrestricted. Finally, we estimate \( q_{21} \), which is the main parameter of interest. Under the assumption that the first shock is the spending shock, this coefficient measures the state-dependent impact of exogenous increases in government spending on output. To summarize, we impose the following restrictions:

\[
B = \begin{bmatrix}
** & 0 \\
** & 0 \\
*** & \end{bmatrix}, \quad Q = \begin{bmatrix}
* & 0 & 0 \\
0 & * & 0 \\
0 & 0 & * \\
\end{bmatrix}, \quad \Lambda_2 = \begin{bmatrix}
1 & 0 & 0 \\
0 & 1 & 0 \\
0 & 0 & *
\end{bmatrix},
\]

(6)

where asterisks denote unrestricted elements.

While \textit{a priori} we do not know the ordering of the spending and the output shock, we can determine it after estimation based on the decomposition of the forecast error variances. It turns out that the ordering of the shocks is the same as before. The
government spending shock is ordered first, explaining more than 90% percent of the variability of government spending in both states and at all horizons. The technology and the monetary policy shock follow with variance contributions to GDP and the T-Bill rate of, respectively, 60% on average across states and horizons.

The solid lines in Figure 9 show the cumulative government spending multiplier in both regimes for the decomposition (5) with the restrictions (6), together with 90% confidence bands. In the low volatility regime (left panel), the multiplier is 2.8 upon impact and increases slightly in the subsequent quarters. It reaches a maximum one year after the shock occurred. Then, it gradually declines but is still significantly larger than one after four years. At the end of the response horizon, it falls to 1.3. In sharp contrast, the multiplier is substantially smaller in the high volatility regime. It is 1.2 on impact but the effect is not significantly larger than one. Thereafter, it falls below unity. After five years, the value is 0.5. A comparison across regimes shows that the multiplier is twice as large in the low as in the high volatility regime. Moreover, as Figure 10 shows the difference between both multipliers is also statistically significant for all horizons.

Figure 9: State-dependent government spending multipliers. Notes: The figure shows the government spending multiplier in the low volatility regime (left panel) and in the high volatility regime (right panel). The solid line and shaded area correspond to the point estimate and 90% confidence bands, respectively, for decomposition (5) and the dashed line refers to the point estimate implied by a Cholesky decomposition with ordering $y_t = [g_t, x_t, r_t]'$.

To assess the sensitivity of the state-dependency to the identification scheme, the dashed lines in Figure 9 show regime-specific multipliers using a Cholesky decomposition in both states with government spending ordered first, followed by output and the T-Bill rate. Relative to the decomposition (5) with (6), this identification strategy implies that the change in variances is shifted entirely to the main diagonal of $Q$, that the elasticity of
government spending to output is zero in both regimes, and that the response of monetary policy is state-dependent. Again, the multiplier is larger in the high than in the low volatility regime and this difference is statistically significant (we omit the respective error bands for visibility). Moreover, in both states the point estimate is below the one implied by the generalized decomposition, in line with the argument that the timing assumption of Blanchard and Perotti (2002) produces lower multipliers. Finally, the differences between identification schemes are larger in the low volatility regime, reflecting a larger estimated absolute sensitivity of government spending to output in this state.

Figure 10: Difference between state-dependent government spending multipliers. Notes: The figure shows the difference between the government spending multiplier in the low volatility regime (left panel of Figure 9) and in the high volatility regime (right panel of Figure 9). The solid line and shaded area correspond to the point estimate and 90% confidence bands, respectively, for decomposition (5) and the dashed line refers to the difference between the state-dependent point estimates implied by a Cholesky decomposition with ordering \( y_t = [g_t, x_t, r_t]' \) as displayed in Figure 9.

Summarizing, we find evidence for state-dependent effects of fiscal policy. The government spending multiplier is significantly larger when uncertainty is low, consistent with the theoretical predictions of, among others, Basu and Bundick (2017) and Bloom et al. (2018). For the general decomposition, the state-dependent multipliers span the average multiplier implied by the unrestricted MS model (see Table 6). On impact, for example, they are 2.8 and 1.2 in state 1 and 2, respectively, and 2.2 on average. For the recursive ordering, they are 1.6 and 0.9 on impact in state 1 and 2, respectively, and 1.4 on average.

The findings also rationalize that we find larger average multipliers in the heteroskedastic than in the linear models. The average multipliers in the first year, for example, are 1.4 and 0.9 for the MS models, while they are 1.0 and 0.8 for the linear models. These
differences reflect differences in the weighting of the state-dependent effects. The smaller multiplier estimates for the high uncertainty regime obtain less weight in the MS models than in the linear models, producing larger average estimates.

These findings contain important policy implications. Combined with the smoothed state probabilities (see Figure 2), they imply that fiscal policy tended to be less effective in the first part of the sample because there was a larger probability of being in the high volatility regime. In contrast, fiscal policy was more effective during the Great Moderation and the years that followed the Great Recession. This might be driven by the more stable policy environment in the second part of the sample. By following a well-defined and transparently communicated rule-based behavior, such a policy framework can reduce uncertainty and simplify expectation formation by private agents. When future policy interventions and output fluctuations are easier to predict, the precautionary savings motive of households is reduced and firms’ region of inaction shrinks.

6 Conclusion

We estimate the macroeconomic effects of government spending shocks in the US. We identify the shocks through an agnostic identification scheme that exploits time-varying volatility in quarterly post-WWII data modeled through a Markov switching approach.

We find that the average output effects of exogenous spending increases are positive and that the multiplier is significantly and persistently above one. This is mostly due to a crowding-in of private consumption and non-residential investment. The estimated average multiplier is larger than multipliers estimated based on timing-restrictions or exogenous increases in military spending. There are three main reasons that explain these differences. First, we find that government spending responds contemporaneously countercyclical to exogenous changes in output. As the size of the government spending multiplier is negatively affected by the size of this elasticity (Caldara and Kamps, 2017), estimating a negative elasticity implies a larger multiplier. Second, we find that exogenous changes in defense spending produce smaller total government spending multipliers than exogenous changes in total government spending. Third, accounting for the heteroskedasticity in US macroeconomic data implies an efficient weighting of observations. The generalized model puts more weight on observations with low error variance compared to observations with high error variance, producing a different estimate than linear models with constant variance and weighting over time.

In the second part of the paper, we exploit that the Markov switching model naturally lends itself to an evaluation of the state-dependent effects of government spending shocks. We show that the multiplier is considerably smaller in high than in low volatility regimes, consistent with an increased option value of waiting that reduces the effectiveness of fiscal

From a policy perspective, our findings imply that government spending is an effective tool for stabilizing the economy. However, they also contain a grain of salt as they suggest that the efficacy of such interventions is lower when both policy and economic uncertainty is high. This indicates that a transparent communication along with rules-based monetary and fiscal policy can lay the ground for successful macroeconomic stabilization.

References


