BUSINESS SCHOOL
Te Kura Pakihi

ISSN 1178-2293 (Online)

University of Otago
Economics Discussion Papers
No. 2302

MARCH 2023

Government Spending Multipliers: Is there a Difference Between Government Consumption and Investment Purchases?

Alfred A. Haug†
Department of Economics, University of Otago, PO Box 56, Dunedin 9054, New Zealand
Email: alfred.haug@otago.ac.nz

and

Anna Sznajderska
Warsaw School of Economics, al. Niepodległości 162, 02-554 Warsaw, Poland.
Email: asznajd@sgh.waw.pl

Address for correspondence:

Alfred Haug
Department of Economics
University of Otago
PO Box 56
Dunedin
NEW ZEALAND
Email: alfred.haug@otago.ac.nz

Government Spending Multipliers: Is there a Difference Between Government Consumption and Investment Purchases?

Alfred A. Haug[†]
Department of Economics, University of Otago, PO Box 56, Dunedin 9054, New Zealand
Email: alfred.haug@otago.ac.nz

and

Anna Sznajderska
Warsaw School of Economics, al. Niepodległości 162, 02-554 Warsaw, Poland.
Email: asznajd@sgh.waw.pl

Abstract

This paper empirically studies the U.S. multiplier effects of government investment, government consumption and total government purchases on output. We explore dependencies of the multipliers on states of the economy, measured in different ways. Using local projections with instrumental variables, we find that a model without state-dependencies and using total government spending (instead of its components) provides the best fit to post-WWII data. These results are robust to various alternative specifications. We account for the COVID-19 period with a pandemic stringency index and for monetary policy shocks with a shadow interest rate. The government spending multiplier is approximately 0.5.

JEL Classifications: E62, E63, H50.

Keywords: Government investment and government consumption multipliers; COVID-19; local projections.

[†]Corresponding author.

1 Introduction

On a theoretical level, neoclassical- and Keynesian-type macroeconomic models generally predict an increase in output for a positive government spending shock. However, the size of the multiplier critically depends on how the interest rate responds to the shock and how much crowding out of private investment and consumption takes place.[1] There is no agreement in the literature on the size of the government spending multiplier (Ramey 2019).

The recent empirical literature on government spending multipliers studies whether multipliers differ across states of the business cycle, in periods of interest rates near their zero lower bound, and during periods of relatively high private or government debt (e.g., Auerbach and Gorodnichenko 2012, 2013; Bernardini and Peersman 2018; Fazzari et al. 2015, 2021; and Ramey and Zubairy 2018). In other words, multipliers may be state-dependent and hence nonlinear. These studies focus on total government spending and do not consider separately its components, government investment and government consumption purchases.[2] An exception is Auerbach and Gorodnichenko (2012, 2013), but they do not test for the equivalence of the government investment and consumption multipliers, nor do they test whether they are linear or nonlinear multipliers. Ramey and Zubairy (2017, 2018) re-examine Auerbach and Gorodnichenko's results using local projection methods and consistently defined cumulative multipliers and find very different results, broadly in agreement with their own findings.[3] Yet, Ramey and Zubairy (2017, 2018) analyse only total government spending and not its components.

In contrast, Boehm (2020) estimates and reports separately government investment and government consumption multipliers that are statistically significantly different from each other[4] and argues that a model that combines the two government spending components into total government spending is a misspecified model. Boehm (2020) employs a panel data set with 17 OECD countries, including the U.S., with quarterly observation from 2003 to 2016. However, Bom and Lighthart (2014) present empirical evidence that the productivity of government capital is country (location), time and type-of-investment specific. This implies that an analysis for a single country may produce results that

[1] See, for example, Blanchard and Perotti (2002).
[2] Ben Zeev et al. (2023) extend Ramey and Zubairy's (2018) framework to allow for shocks to total government spending to differ by sign but find no evidence for asymmetric government spending multipliers.
[3] Ramey and Zubairy (2018) study the U.S. government spending multiplier during periods of economic slack with historical data. After a military news shock, the spending multiplier is 0.60 and 0.59 after two years, in times of high and low economic slack, respectively. The difference across states is not statistically significant, i.e., the spending multiplier is linear.
[4] Boehm (2020, p. 88, Table 1) reports two *p*-values for testing the hypothesis that the output multipliers are the same for government consumption and government investment with values of 0.061 at the fourth quarter horizon and of 0.057 at the eighth quarter horizon. Hence, only a 10% level of significance supports rejecting the null hypothesis of equal multipliers but not a 5% level.

differ from those for a panel. Furthermore, Boehm (2020) does not consider nonlinear multipliers and Boehm's linear multipliers represent hence averages over the business cycle. Fazzari et al. (2015) find that an increase in total government spending crowds out private investment in economic expansions but not in recessions. A linear multiplier may show instead a negative response because it is approximately a weighted average of the responses across the business cycle. Therefore, it is crucial for model building and economic policy guidance to explore whether multipliers are state-dependent.

This paper contributes to the literature by extending the analysis of government spending multipliers to its components, government investment and government consumption for the U.S. economy. The analysis applies state-of-the-art local projection methods with instrumental variable estimation in order to directly compute cumulative multipliers. It formally tests whether multipliers for U.S. government investment and government consumption are the same, and in addition whether the government investment and consumption multipliers depend on unemployment levels, recessions or private debt relative to GDP. Further, we include the COVID-19 period and assess the performance of a pandemic stringency index. Section 2 briefly describes the econometric methodology and Section 3 present the results, followed by a sensitivity analysis in Section 4. Section 5 concludes.

2 Methodology

Following the recent empirical literature on government multipliers, we apply the local projection (LP) method of Jordà's (2005) to estimate impulse responses that represent also the multipliers in our specifications below. Montiel Olea and Plagborg-Møller (2021) show that inference in lag-augmented linear LP-based IRFs is valid with stationary and nonstationary data. We use lag-augmented LP regressions below (the control variables always include lags of the dependent variable). Furthermore, Plagborg-Møller and Wolf (2021) prove the equivalence of impulse responses from VARs and LPs for unrestricted lag structures, even when the data generating process is nonlinear. They also show the validity of LP-IV estimation for impulse responses, regardless of whether or not the shock of interest is non-invertible or the instrumental variable has measurement error (that is unrelated to the shock of interest). However, for finite samples with restricted lags they (p. 956) give the following advice: "At short impulse response horizons, the two methods are likely to approximately agree if the same lag length is used for both methods. However, with finite lag lengths, the two methods may give substantially different results at long horizons." This points to the importance of including sufficient lags in LP estimations. In addition, Montiel Olea and Plagborg-Møller (2021) note that for robust inference over longer horizons, LP regressions should be augmented with the same number of lags for all variables. Montiel Olea and Plagborg-Møller (2021) and Li et al. (2022) prove that standard confidence intervals based on lag-augmented LPs have uniformly correct asymptotic coverage over a wide range of horizons. However, in finite sample, a trade-off between traditional VARs and LPs can emerge between a higher bias of coefficient estimates in VARs and a higher variance of coefficient estimates in LPs (Li et al. 2022). An example is provided by Lyu and Noh (2022) for inference on U.S. (total) government spending multipliers with a Markov-switching VAR and alternatively with LPs, with the latter having larger confidence intervals (meaning less precise, i.e., more inefficient, estimation of intervals) and making it more difficult to detect statistical significance. Due to these concerns, we will check the sensitivity of our results to lag lengths in our LP regressions.

LPs are used with instrumental variables (IVs) to estimate the cumulative multipliers and impulse responses directly (e.g., Bernardini and Peersman 2018; and Ramey and Zubairy 2018). Furthermore, as in Ramey and Zubairy (2018), we apply a transformation proposed by Gordon and Krenn (2010) with variables specified in levels, except for real GDP and real government spending (or its components), which are divided by the trend of real GDP.[5] The Gordon-Krenn transformation leads to multipliers that give the response of real GDP in dollars for a one dollar shock to government investment, government consumption or total government purchases, i.e., dollar-for-dollar multipliers. We apply the Hodrick-Prescott (HP; Hodrick and Prescott 1997, 2020) filter to extract a stochastic trend from real GDP with the smoothing parameter λ set equal to 1,600, as usual with quarterly data.

For each horizon *h* the following nonlinear model is specified:

$$\tilde{y}_{t+h} = I_{t-1}\left[\alpha_{A,h} + \psi_{A,h}(L)z_{t-1} + \beta_{A,h}\widetilde{shock}_t\right] + (1 - I_{t-1})\left[\alpha_{B,h} + \psi_{B,h}(L)z_{t-1} + \beta_{B,h}\widetilde{shock}_t\right] + \varepsilon_{t+h}, \quad h = 0, 1, \dots, 12, \qquad (1)$$

where $\tilde{y}_{t+h}$ is real GDP in period *t* divided by its stochastic trend. The indicator variable, I_{t-1}, for the state of the economy at time *t-1* is set to 1 for periods of high unemployment or recessions, labelled state *A*, and to 0 for periods of low unemployment or expansions, labelled state *B*. Alternatively, "state" refers to high (with a value of 1) or low (with a value of 0) private domestic non-financial-sector debt relative to GDP, depending on the empirical specification being considered. To avoid contemporaneous feedback from government policy to the state, the state is determined by the state in the period before the shock hits. $\widetilde{Shock}_t$ is a government investment, or alternatively, a government consumption or total government purchases shock, divided by the stochastic trend of real GDP. The shock is uncorrelated with the error term ε_{t+h}, and $\beta_{A,h}$ and $\beta_{B,h}$ are the non-cumulative responses of real GDP in states *A* and *B*, respectively, at a specific horizon *h* after the shock in period *t*. The

[5] An alternative transformation used by Bernardini and Peersman (2018) is instead to scale the changes in each of real GDP and government spending (or its components) by y_{t-1}. We applied this method to the baseline model and results are essentially the same.

parameter α is a constant, $\psi(L)$ is a lag polynomial, and z_{t-1} is a vector of lagged control variables that includes lags on $\tilde{y}_t$ and government investment and consumption. In all regressions, four lags on all control variables are included, following Ramey and Zubairy (2018), among others.

2.1 Cumulative Multipliers

As in Ramey and Zubairy (2018), the cumulative multiplier is estimated directly with the following equation and the LP-IV method:

$$\sum_{j=0}^{h} \tilde{y}_{t+j} = I_{t-1} \left[\gamma_{A,h} + \phi_{A,h}(L) z_{t-1} + m_{A,h} \sum_{j=0}^{h} \tilde{g}_{t+j} \right] + (1 - I_{t-1}) \left[\gamma_{B,h} + \phi_{B,h}(L) z_{t-1} + m_{B,h} \sum_{j=0}^{h} \tilde{g}_{t+j} \right] + \omega_{t+h},$$
$$h = 0, 1, \dots, 12, \qquad (2)$$

where $I_{t-1}\ \widetilde{shock}_t$ and $(1 - I_{t-1})\widetilde{shock}_t$ are used as instruments for the cumulative government investment to trend GDP ratio, or alternatively, the government consumption or total government purchases to trend GDP ratio, $\sum_{j=0}^{h} \tilde{g}_{t+j}$, in the corresponding state. The estimated cumulative spending multipliers at horizon *h* are the estimates of $m_{A,h}$ and $m_{B,h}$ for state *A* and *B*, respectively.

The shock and the government spending variable (or one of its components) are both allowed to have measurement error, as long as their measurement errors are uncorrelated. The two-stage least-squares econometric software package "Stata" command *ivreg* (Baum et al. 2010) is used for estimation. Autocorrelation in the presence of IVs is tested for with the *ivactest* test (Baum et al. 2007). This test reveals serial correlation and we therefore run all IV regressions with the Newey-West correction with automatic bandwidth selection, *bw(auto)* and *robust,* to ensure heteroskedasticity and autocorrelation consistent (HAC) standard errors.

2.2 Government Investment and Government Consumption Shock Identification

Government spending shocks, referred to as Blanchard-Perotti shocks by Bernardini and Peersman (2018) and Ramey and Zubairy (2018), are identified by assuming that government purchases, and its components likewise, do not respond contemporaneously to other structural shocks. This is a reasonable assumption for the identification of structural government spending shocks, including for its components, government investment and government consumption purchases. It is assumed that government spending policy is backward-looking and evolves as:

$$\tilde{g}_t = \psi(L) z_{t-1} + \widetilde{shock}_t, \tag{3}$$

where $\tilde{g}_t$ is determined by the same set of lagged control variables used in equation (1) above, z_{t-1}, and an orthogonal shock ($\widetilde{shock}_t$) of autonomous changes to $\tilde{g}_t$. Depending on the model specification being studied, $\tilde{g}_t$ stands for either government investment, government consumption or total government purchases.

2.2.1 Weak Instruments and Tests

Weak IVs may lead to biased estimates (Andrews et al. 2019). We calculate, for the first stage estimates of the two-stage-least-squares procedure of the LP-IV method, the Kleibergen-Paap rk F-statistic (Kleibergen and Paap 2006) for each state and horizon of equation (2). An F-statistics value below 10 is generally used as indicating a weak instrument (Staiger and Stock 1997; Stock and Yogo 2005). Montiel Olea and Pflueger (2013) propose a stricter threshold of 23.11 at a 5% significance level when first stage regression errors are heteroskedastic and serially correlated.[6]

We test whether the cumulative multipliers are significantly different across the states *A* and *B*. The null hypothesis is that the cumulative multipliers are the same in both states, at a given horizon. If the null hypothesis cannot be rejected, the model is linear and multipliers are not affected by unemployment levels, recessions, or private debt levels (depending on the definition of the threshold for the indicator variable):

$$\sum_{j=0}^{h} \tilde{y}_{t+j} = \gamma_h + \phi_h(L) z_{t-1} + m_h \sum_{j=0}^{h} \tilde{g}_{t+j} + \omega_{t+h}, \quad h = 0, 1, \ldots, 12 \tag{4}$$

If the null hypothesis is rejected, multipliers are dependent on the state. The tables in our paper report the HAC-based *p*-values when instruments are not weak, and Anderson and Rubin's (AR) (1949) test *p*-values when instruments are weak (i.e., when the rk F-statistic value is below 10). This is because the AR test is robust to weak instruments (but possibly less powerful than the HAC-based test), whereas the HAC-based test is not robust to weak instruments.

3 Results

Quarterly U.S. data used in the analysis are described in Appendix A. The sample period covered is from the first quarter of 1947 to the second quarter of 2022. The start date is due to data availability and the end date reflects the latest data available when this research was started. In particular, the government spending components, consumption and investment, are not available prior to 1947.

---

[6] It is the value for a threshold of a 5% critical value for testing the null hypothesis that the two-stage least squares bias exceeds 10% of the ordinary least squares bias. The 10% threshold value is 19.75.

In regards to the COVID-19 period, we follow Ng (2021) and treat COVID as a model-exogenous "health" shock that is not captured by the other shocks in the model. Ng controls for COVID-19 with indicators and finds that the macroeconomic VAR-based impulse response functions are very similar to those of the pre-COVID period. This shows that indicator data for the pandemic are able to control for outliers adequately. Similar to Ng (2021), we include a composite COVID-19 stringency index for the U.S. in all full-sample regressions to capture the effects of the pandemic. It has non-zero values from the first quarter of 2020 onwards.

The choice of the other variables for inclusion in our models is guided by empirical specifications employed in related recent studies on fiscal multipliers in the pre-COVID-19 period. We include the same variables that Blanchard and Perotti (2002) use in their seminal study but augment the set of variables with the interest rate (shadow rate) and inflation rate to account for monetary policy, aside from using the components of government spending instead of the total. Rossi and Zubairy (2011), among others, advocate studying monetary and fiscal policy shocks together to avoid attributing output and other economic effects to the wrong source.

3.1 Baseline Model: States Defined by an Unemployment Threshold

The model in equation (2) is estimated with the state of the economy defined as high unemployment when the unemployment rate is at or above 6.5%. The state of low unemployment is defined as quarters when the unemployment rate is below 6.5%.[7] This is the same threshold used by Ramey and Zubairy (2018), among others, based on U.S. monetary policy announcements during Ben Bernanke's tenure as chair of the Board of Governors of the Federal Reserve System (2006 to 2014). They argue that with alternative business cycle (recession and expansion) measures, such as an output gap based on the moving average growth rate of GDP, only half of the quarters classified as recessions in their sample show high unemployment (at or above 6.5%) and they question its value as an indicator of the state of economic slack.

The baseline model includes the following variables:

- real GDP;
- real government investment or, alternatively, real government consumption purchases;
- real government current net tax receipts;
- inflation rate (CPI-based);
- nominal shadow federal funds rate;

[7] There are 178 quarters of high unemployment and 124 quarters of low unemployment among the 298 quarterly observations (after accounting for lags used in the regressions).

- COVID-19 stringency index.

This set of variables turns out to lead to non-predictability of the Blanchard-Perotti-type government investment and consumption shocks used in our analysis.[8] A model with less variables included in the regressions leads to shocks being predicable.[9]

Table 1 presents the estimates for the cumulative government investment multiplier for the linear model in equation (4) and the non-linear model in equation (2), for even-numbered horizons to conserve space. It also reports the rk F-statistics for weak instruments and the *p*-values for the test of the null hypothesis that the multipliers are the same in states of low and high unemployment. The null hypothesis cannot be rejected for all horizons. Hence, there is no empirical support for the effects of government investment shocks on real output being state-dependent. In other words, the linear model with no state-dependency is empirically supported. Furthermore, the results are not compromised by weak instruments because F-statistics in the linear case are all well above any thresholds. In the non-linear specification there is one weak IV case with an F-statistic below the threshold value of 10 for the state of high unemployment at horizon 12. However, the weak-instrument-robust AR test statistic *p*-value for the hypothesis of equal multipliers across states is well above usual critical values and the null hypothesis cannot be rejected.

The linear model produces a government investment multiplier that peaks at horizon h=1 with a value of 1.11 and then steadily falls to 0.39 at h=5, becoming a case of borderline insignificance at the 5% level from h=5 to h=7. At the 10% level, the linear multipliers become insignificant after h=7. Figure 1 depicts the cumulative investment spending multipliers with 95% confidence intervals.

Table 2 shows corresponding results for the government consumption purchases multipliers. Based on the calculated *p*-values, the null hypothesis that consumption multipliers are the same across states of low and high unemployment cannot be rejected. Again, the linear model is backed by the data. Instruments are not weak, except for the state of high unemployment at h=12, where the AR-based test for equal multipliers across states has a *p*-value of 0.30. The linear specification leads to a peak multiplier of 0.40 at h=3. However, none of the linear multipliers is statistically significant at usual levels. Figure 2 displays the linear government consumption multipliers along with 95% confidence intervals.

.

[8] The shock variables estimated from equation (3) are regressed for all horizons $t+h$, h=1, 2, … 12, on the information set of the empirical model at date *t-1*, i.e., on four lags of the control variables (z_t), following the approach in Favero and Giavazzi (2012). Non-predictability (orthogonality) implies that the lagged variables are not statistically significant. Results are available from the authors on request. See also Ramey (2011) and Bernardini and Peersman (2018, fn. 14, p. 492).

[9] On the other hand, replacing the CPI-based inflation rate with an inflation rate based on the GDP-deflator produces very similar results and no qualitative differences.

3.2 States Defined by a Moving Average of GDP Growth

The model estimated is the same as in Section 3.1, except that states *A* and *B* are defined based on recessions and expansions over the business cycle, instead of an unemployment-rate threshold. We follow Auerbach and Gorodnichenko (2012) and define a smooth transition threshold variable ranging from 0 (sufficiently strong expansion) to 1 (sufficiently deep recession) based on a normalized seven-quarter centred moving average of GDP growth.[10] Moreover, this variable highly correlates with the NBER official U.S. business cycle expansions and contractions (Ramey and Zubairy 2017, Figure 4).

Table 3 produces the results for the continuous recession indictor applied to equation (2). The multipliers for expansions and recession are generally lower, reaching negative values after *h*=2 for expansions and after *h*=4 for recessions, when compared to states of low and high unemployment in Table 1. Nevertheless, the null hypothesis of multipliers being equal in expansions and recessions cannot be rejected with all *p*-values at or above 0.39 across horizons.

Government consumption expenditure shocks in Table 4 reveal cumulative multipliers that are commonly somewhat smaller in expansions compared to low unemployment states in Table 2. On the other hand, they are larger in recessions than in high unemployment states in Table 2, peaking at 0.61 at *h*=2 and falling steadily to -0.04 at *h*=12 in Table 4. Regardless, the null hypothesis that multipliers are the same across business cycle states cannot be rejected, as was the case in Table 2. The p-values in Table 4 are all at or above 0.46 and instruments are not weak.

The inference with Auerbach and Gorodnichenko’s (2012) alternative definition of states *A* and *B* leads to the same conclusion as in Section 3.1. The linear model specification is supported again. Government investment and government consumption multipliers are not state-dependent.

.

3.3 States Defined by the Relative Level of Private Debt

Bernardini and Peersman (2018) point out that several theoretical and empirical studies have shown that private indebtedness amplifies business cycles, though government spending multiplier dependence on private indebtedness had not been considered previously at that point in time and theirs is the first study to fill this gap. We follow Bernardini and Peersman’s definition of private debt in their baseline model and use the ratio of nominal private domestic non-financial-sector debt to nominal GDP. High private debt is defined as periods when the private debt to GDP ratio is above its stochastic trend, noted as state *A* with an indicator value of 1, for at least two consecutive quarters. Otherwise, the state is defined as a low debt state *B* with an indicator value of zero. Keeping in line with Bernardini

[10] The moving average growth rate has a mean of 0.76 and a standard deviation of 0.50 in our sample. We use approximate values of 0.8 and 0.5, respectively, which are, according to Ramey and Zubairy (2018), the values used by Auerbach and Gorodnichenko (2012).

and Peersman (2018), we extract a very smooth HP trend for the private debt to GDP ratio by setting $\lambda=10^6$.

Table 5 records results for the government investment multiplies in low and high private debt states. In low debt states, the multiplier peaks at 0.85 at h=1 and then fluctuates between 0.73 and 0.37 over the other horizons. In the high states, the multiplier reaches a maximum of 1.30 at h=1 and then falls across horizons to -0.02 at h=12. The evidence points to strong instruments for all horizons across the two states. Next, we test again the crucial null hypothesis that multipliers do not differ across states. It cannot be rejected with p-values at or above 0.20 for each and every horizon. This result confirms the relevance of the linear specification for the government investment multiplier.

In Table 6, we supply the government consumption multiplier results. Low private debt states furnish lower multiplier values, compared to the low unemployment state in Table 2. In contrast, higher private debt states reveal larger multipliers than the high unemployment state in Table 2, ranging from 0.33 at h=2 to 0.70 at h=8. Instruments are not weak at all horizons except for h=12 for the low debt state. The hypothesis of equal multipliers across low and high private debt states cannot be rejected for the government consumption multipliers based on p-values being all at or above 0.22 and using the AR-based test at h=12.

.

3.4 Testing the Equivalence of Government Investment and Government Consumption Purchases Multipliers

The analysis of government spending multipliers in the literature generally studies the effects of government spending shocks in the short run over the business cycle. For the post-WWII period a typical horizon for the effects of shocks is 12 quarters. Furthermore, a shock is formulated as a one-time temporary impulse and it is not a sustained change over several quarters.[11] Boehm (2020) argues that temporary government investment shocks crowd out private investment, whereas government consumption shocks do not. Boehm finds in a panel study empirical support for short-run government consumption multipliers that are larger than short-run government investment multipliers, with the former multiplier value approximately 0.8 and the latter near zero.

We proceed to testing the null hypothesis that the government investment and government consumption multipliers are the same in a given state A or B for every horizon after the initial shock over the full sample. The non-linear models with thresholds based on the unemployment rate, the GDP growth rate and the relative private debt levels were not supported by the data, as shown in the previous

[11] Baxter and King (1993) find in a quantitative equilibrium model that permanent changes in government purchases have larger multipliers than temporary changes, and long-run output multipliers can be larger than 1.

three sections. Therefore, we give in Table 7 (in the second column) p-values of tests for the null hypothesis for the linear model, which is, of course, the same for all three previous specifications in Section 3.1 to 3.3. The null hypothesis that government investment and consumption multiplier are the same cannot be rejected at every horizon, with all p-values above 0.16. Therefore, in contrast to Boehm's (2020) panel study and in agreement with studies employing total government spending such as Ramey and Zubairy (2018) and Bernardini and Peersman (2018), the short-run temporary total government spending multiplier adequately captures the effects on output and there is no need to disaggregate government spending into its components.

4 Sensitivity Analysis

We explore various alternative specifications of the baseline model in Section 3.1 to assess the robustness of the empirical findings so far. First, we exclude the pandemic period for the baseline model. Then, we present results for total government spending for the three alternative non-linear model specifications. As all evidence supports a linear specification, we next report linear results for adding government debt to the baseline model as an additional variable and for using a sixth-order polynomial exponential trend of real GDP for Gordon-Krenn transformations instead of an HP trend. Likewise, we explore sensitivities to including military news among the control variables and to extending the lag length from four to 12 lags for z_t.

4.1 Results Without the Pandemic Quarters

The period without the recent pandemic ranges from 1947Q1 to 2019Q4. Table 8 exhibits results for the linear and non-linear models for government investment multipliers in the baseline model with the unemployment threshold to select states *A* and *B*. The null hypothesis that the multipliers are the same across states at each horizon cannot be rejected at usual significance levels, with p-values similar to those in Table 1, where the pandemic period is included along with its stringency index. Also, the linear model in Table 8 reveals very similar multipliers to those in Table 1.

The same pattern emerges when we compare the results for government consumption multipliers without the pandemic period in Table 9 to those with it in Table 2. Again, the linear model is supported by the data and linear multipliers are very similar across the two tables. What remains is to test whether the linear government investment multiplier is the same as the government consumption multiplier for every horizon. This hypothesis test furnishes p-values in the range from 0.18 to 0.40 in the last column of Table 7.

We conclude this section by noting that the COVID-19 stringency index adequately captures the economic effects of the pandemic. Linear multipliers are similar for the sample with and without the pandemic period. Hence, the pandemic period seems not to have led to structural breaks in the underlying economic effects of government spending on output.

4.2 Results for Total Government Spending: The Three Non-Linear Models

Our empirical results do not lend support to separating total government spending into its components. The multipliers for government investment and government consumption expenditures are not statistically significantly different form each other at every horizon considered. Therefore, we add results for total government spending and concentrate henceforth on total government spending instead of its components.

The first model is the baseline model with the unemployment threshold and total government spending instead of its investment and consumption components. Table B.1 records results. One would expect that multipliers are not state-dependent for total government spending because its components are not. Indeed, the null hypothesis of equal multipliers across states of low and high unemployment cannot be rejected, given that *p*-values in Table B.1 are above 0.38 for all horizons. The linear model produces multipliers that peak at 0.50 at h=1 and then steadily fall to 0.02 at h=12. The effects of the shocks on output are statistically significant at the 10% level up to, and including, h=4.

Table B.2 shows estimates for total government spending for the non-linear model with Auerbach and Gorodnichenko's (2018) continuous recession indicator variable. Table B.3 records regression output for the non-linear model with the private debt-to-GDP trend as the threshold for low and high private debt. Both alternative thresholds lead to the same conclusion that the linear model (Table B.1) is the model that is empirically relevant. Our results are consistent with those of Ramey and Zubairy (2018), favouring a linear specification where government spending multipliers are not state-dependent (Figure B.1).

4.3 Additional Sensitivity Analysis

We include the government debt-to-GDP ratio as an additional control variable in the linear baseline model in equation (4) for total government spending. Bernardini and Peersman (2018) discuss how government debt may affect government spending multipliers. Higher government debt implies higher expected future tax rates (when tax distortions are convex), resulting in a negative wealth effect on private consumption. Thus, government spending shocks could have a positive effect on private consumption when government debt is low, but not when government debt is high relative to GDP. On the other hand, if Ricardian equivalence holds the way government spending is financed is irrelevant, be it current taxation or issuing debt (future taxation) and only the path of government spending matters to the private sector. Bernardini and Peersman (2018) do not find an unambiguous effect of government debt on the government spending multiplier. In our analysis, the government debt-to-GDP ratio does not enter statistically significantly in the regression equations. The *p*-values for the null hypothesis that all coefficients of the lags of the debt variable are jointly insignificant are above 0.24 at all horizons.

Next, we consider a sixth-order polynomial trend for the natural logarithm of real GDP. It is applied to normalise real GDP and total government spending in the Gordon-Krenn transformation when calculating dollar-for-dollar multipliers. Ramey and Zubairy (2018) use this approach instead of an HP filter that we used above. It is mostly motivated by the poor performance of the HP filter in the pre-WWII period (Gordon and Krenn 2010). Ramey and Zubairy's (2018) sample covers the period from 1889Q1 to 2015Q4. Still, it is of interest to check how the two methods of normalisation compare for the post-WWII period.

Table B.4, in columns two and three, presents the linear multipliers and the rk F-statistics, showing that instruments are not weak, when applying a polynomial trend for the Gordon-Krenn transformations. Results are again for the baseline model with the unemployment threshold and total government spending. Multipliers are similar in magnitude to those with the HP trend (in Table B.1), though somewhat larger in magnitude with a peak value of 0.60 at h=1 compared to 0.50 at h=1 for the HP trend. Furthermore, other qualitative results, in particular for the state-dependency and the equivalence of government investment and consumption multipliers, are the same as with the HP trend.

The government spending shocks are assumed to be surprise shocks and should not be correlated with other shocks. One way to check whether other shocks affect our calculated multipliers is to include them among the set of control variables. We add the narrative military spending news shocks from Ramey and Zubairy (2018) to the baseline model, with updates beyond 2015Q4, when their sample ends, assuming no major military shocks after 2015Q4. Table B.4, in columns four and five, shows results. Qualitative results are the same as in the specification without the military news variable and the linear model emerges again as the favored specification. Multipliers differ little from those in Table B.1 where the military news variable is excluded. The peak value in both cases is 0.50 at h=1 and the difference in multipliers is not more than 0.01 at any horizon, and they are the same at horizons 1, 3, 4, and 8.

Last, we address the issue of LP-IV consistency and efficiency for the impulse response estimates, particularly the potential inefficiency of the confidence interval estimates that matter for our hypothesis tests. Large standard errors of the coefficient estimates may lead to being unable to rejects the various null hypothesis that we tested. Plagborg-Møller and Wolf (2021, pp. 974-975) confirm that for finite samples with restricted-lag LPs with p lags for the controls and VARs with p lags "agree approximately at impulse response horizons $h \leq p$." Hence, such "VARs will tend to agree at short and medium-long horizons" with LPs that also control for a rich set of lags. Accordingly, we explore whether our results hold up for the baseline model when we increase the lag length from four to 12 lags. Table B.5 furnishes the hypothesis tests for government multipliers being the same for government investment and government consumption, as well as for the total government spending multiplier being the same across low and high unemployment states. The results for the hypothesis tests are unchanged when compared to the specification with only four lags. What is more, the linear total government spending multiplier LP-IV estimates in Table B.5 with 12 lags are close to those in Table B.1 with four lags, with only slightly smaller values (by 0.06 or less). This demonstrates that our empirical findings are robust to adding lags for the control variables.

5 Conclusion

This paper studies the effects of the components of government spending, government investment and government consumption purchases, on output with quarterly U.S. data over the period 1947Q1 to 2022:Q2. We explore whether the government investment and government consumption multipliers depend on the state of the economy, i.e., on unemployment levels, recessions or private debt levels. We find no empirical evidence for state dependency. Next, we explore whether the government investment multiplier is different from the government consumption multiplier. The two multipliers show no statistically significant differences at all horizons.

The sample period includes the COVID-19 pandemic and all full sample specifications incorporate a pandemic stringency index. The index seems to adequately capture the pandemic's economic effects. The pre-pandemic model produces very similar results to the full-sample model.

We apply local-projection methods with instrumental variables to directly estimate cumulative multipliers. This method allows for measurement errors in the fiscal shocks that drive the multipliers. In our regressions we account for monetary policy shocks by including inflation and the Wu and Xia (2016) federal funds shadow rate. The latter is capturing the stance of monetary policy when interest rates are near their zero lower bound. Thus, we endeavour to avoid attributing monetary policy shocks to fiscal shocks and consequently to fiscal multipliers. Also, we carry out an extensive sensitivity analysis and conclude that our empirical findings are robust in various directions. Our analysis validates using linear government spending multiplier model specifications that do not depend on the state of the economy and are based on total government spending, not its components. These results are broadly consistent with the empirical evidence provided by Ramey and Zubairy (2018) for a longer sample period.

The linear output multipliers for total government spending are all below 1 across various alternative specifications. These multipliers usually decline as the horizon increases and peak generally around a value of 0.5 in the quarter following the shock. But, they are typically statistically significant only for the first four quarters.

Our empirical results should be useful for macroeconomic model building and economic policy alike. A note of caution is that the multipliers in our paper are based on one-time, temporary and short-run shocks to government spending. Long-run and sustained shocks to government spending may well have different effects on output and multipliers. In order to investigate long-run effects of government spending, economic growth models would seem an appropriate framework.

References

Anderson, T. W., and Rubin, H. (1949). Estimation of the Parameters of a Single Equation in a Complete System of Stochastic Equations. *Annals of Mathematical Statistics, 20*, 46-63.

Andrews, I., Stock, J. H., and Sun, L. (2019). Weak Instruments in Instrumental Variables Regression: Theory and Practice. *Annual Review of Economics, 11*, 727-753.

Auerbach, A. J., and Gorodnichenko, Y. (2013). Corrigendum: Measuring the Output Responses to Fiscal Policy. *American Economic Journal: Economic Policy, 5*, 320-322.

Auerbach, A. J., and Gorodnichenko, Y. (2012). Measuring the Output Responses to Fiscal Policy. *American Economic Journal: Economic Policy, 4*, 1-27.

Baum, C. F., Schaffer, M. E., and Stillman, S. (2010). ivreg2: Stata Module for Extended Instrumental Variables/2SLS, GMM and AC/HAC, LIML and K-class Regression. Retrieved from http://ideas.repec.org/c/boc/bocode/s425401.html, last accessed 5 February 2021.

Baum, C. F., Schaffer, M. E., and Stillman, S. (2007). Enhanced Routines for Instrumental Variables/Generalized Method of Moments Estimation and Testing. *Stata Journal, 7*, 465-506.

Baxter, M., and King, R. G. (1993). Fiscal Policy in General Equilibrium. *American Economic Review, 83*, 315-334.

Ben Zeev, N., Ramey, V. A., and Zubairy, S. (2023). Do Government Spending Multipliers Depend on the Sign of the Shock? NBER Working Paper 31015.

Bernardini, M., and Peersman, G. (2018). Private Debt Overhang and the Government Spending Multiplier: Evidence for the United States. *Journal of Applied Econometrics, 33*, 485-508.

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

Boehm, C. E. (2020). Government Consumption and Investment: Does the Composition of Purchases Affect the Multiplier? *Journal of Monetary Economics*, *115*, 80-93.

Bom, P. R. D., and Ligthart, J. E. (2014). What Have We Learned From Three Decades of Research on the Productivity of Public Capital? *Journal of Economic Surveys, 28,* 889-916.

Favero, C., and Giavazzi, F. (2012). Measuring Tax Multipliers: The Narrative Method in Fiscal VARs. *American Economic Journal: Economic Policy, 4*, 69-94.

Fazzari, S. M., Morley, J., and Panovska, I. (2021). When is Discretionary Fiscal Policy Effective? *Studies in Nonlinear Dynamics and Econometrics, 25*, 229-254.

Fazzari, S. M., Morley, J., and Panovska, I. (2015). State-Dependent Effects of Fiscal Policy. *Studies in Nonlinear Dynamics and Econometrics, 19*, 285-315.

Gordon, R. J., and Krenn, R. (2010). The End of the Great Depression: VAR Insight on the Roles of Monetary and Fiscal Policy. NBER Working Paper 16380.

Hodrick, R. J. (2020). An Exploration of Trend-Cycle Decomposition Methodologies in Simulated Data. NBER Working Paper No. 26750.

Hodrick, R. J., and Prescott, E. C. (1997). Postwar U.S. Business Cycles: An Empirical Investigation. *Journal of Money, Credit, and Banking*, *29,* 1-16.

Jordà, Ò. (2005). Estimation and Inference of Impulse Responses by Local Projections. *American Economic Review, 95*, 161-182.

Kleibergen, F., and Paap, R. (2006). Generalized Reduced Rank Tests Using the Singular Value Decomposition. *Journal of Econometrics*, *133*, 97-126.

Li, D., Plagborg-Møller, M., and Wolf, C. K. (2022). Local Projections vs. VARs: Lessons From Thousands of DGPs. NBER Working Paper 30207.

Lyu, Y., and Noh, E. (2022). Cyclical Variation in U.S. Government Spending Multipliers. *Economic Inquiry, 60,* 831-846.

Montiel Olea, J. L., and Pflueger, C. (2013). A Robust Test for Weak Instruments. *Journal of Business and Economic Statistics, 31*, 358-369.

Montiel Olea, J. L., and Plagborg-Møller, M. (2021). Local Projection Inference is Simpler and More Robust Than You Think. *Econometrica, 89,* 1789-1823.

Ng, S. (2021). Modelling Macroeconomic Variations after COVID-19. NBER Working Paper 29060.

Plagborg-Møller, M., and Wolf, C. K. (2021). Local Projections and VARs Estimate the Same Impulse Responses. *Econometrica*, *89,* 955–980.

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

Ramey, V. A. (2019). Ten Years After the Financial Crisis: What Have We Learned From the Renaissance in Fiscal Research? *Journal of Economic Perspectives, 33*, 89-114.

Ramey, V. A., and Zubairy, S. (2018). Government Spending Multipliers in Good Times and in Bad: Evidence from U.S. Historical Data. *Journal of Political Economy, 126,* 850-901.

Ramey, V. A., and Zubairy, S. (2017). Supplementary Appendix to Government Spending Multipliers in Good Times and in Bad: Evidence from U.S. Historical Data. https://www.journals.uchicago.edu/doi/suppl/10.1086/696277/suppl_file/2014646Appendix.pdf, last accessed 22 November 2022.

Rossi, B., and Zubairy, S. (2011). What Is the Importance of Monetary and Fiscal Shocks in Explaining U.S. Macroeconomic Fluctuations? *Journal of Money, Credit and Banking*, *43*, 1247-1270.

Staiger, D., and Stock, J. H. (1997). Instrumental Variables Regression with Weak Instruments. *Econometrica, 63,* 557-586.

Stock, J. H., and Yogo, M. (2005). Testing for Weak Instruments in Linear IV Regression, in *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg*, D.W. Andrews and J.H. Stock, eds. Cambridge University Press, Cambridge.

Wu, J. C., and Xia, F. D. (2016). Measuring the Macroeconomic Impact of Monetary Policy at the Zero Lower Bound. *Journal of Money, Credit and Banking, 48,* 253-291.

Table 1. Government investment multipliers dependent on the unemployment threshold

Horizon (h)	Linear model (no state-dependency)		State-dependent model				
			Low unemployment		High unemployment		
	Multiplier	F-statistic	Multiplier	F-statistic	Multiplier	F-statistic	*p*-Value for the hypothesis test†
2	0.98**	254.99	1.15	339.48	1.11	57.47	0.98 (HAC)
	(0.34)		(0.45)		(1.23)		
4	0.61**	192.96	1.08	297.83	0.91	20.60	0.91 (HAC)
	(0.24)		(0.43)		(1.39)		
6	0.42**	278.34	1.01	209.19	2.02	12.75	0.31 (HAC)
	(0.21)		(0.46)		(0.82)		
8	0.31	218.96	0.76	126.55	1.46	11.91	0.39 (HAC)
	(0.18)		(0.45)		(0.64)		
10	0.17	172.98	0.46	87.74	1.00	10.14	0.47 (HAC)
	(0.19)		(0.38)		(0.63)		
12	0.05	144.72	0.26	86.41	0.85	8.99	0.52 (AR)
	(0.19)		(0.32)		(0.64)		

Notes: Cumulative multipliers are reported, with their Newey-West (HAC) standard errors given in parentheses. 'F-statistic' is the first stage Kleibergen-Paap rk F-statistic statistic, with weak instruments identified by a value below 10. †The null hypothesis is that the multipliers are the same across states of low and high unemployment. Weak-instrument-robust Anderson-Rubin (AR) *p*-values are reported (instead of HAC-based *p*-values) when the rk F-statistic is below 10 for at least one of the states. For the empirically supported linear model, * denotes statistical significance at the 10% level and ** significance at the 5% level.

Table 2. Government consumption multipliers dependent on the unemployment threshold

Horizon (h)	Linear model (no state-dependency)		State-dependent model				
			Low unemployment		High unemployment		
	Multiplier	F-statistic	Multiplier	F-statistic	Multiplier	F-statistic	*p*-Value for the hypothesis test†
2	0.39	202.78	0.43	360.14	0.17	73.46	0.62 (HAC)
	(0.34)		(0.36)		(0.90)		
4	0.24	114.98	0.40	104.80	-0.39	22.46	0.27 (HAC)
	(0.33)		(0.32)		(0.68)		
6	0.05	68.72	0.23	48.61	-0.72	21.37	0.17 (HAC)
	(0.27)		(0.30)		(0.60)		
8	0.07	46.23	0.23	31.57	-0.62	28.01	0.30 (HAC)
	(0.24)		(0.30)		(0.69)		
10	0.09	37.84	0.19	26.24	-0.65	18.31	0.42 (HAC)
	(0.22)		(0.29)		(0.88)		
12	-0.01	27.99	0.06	19.16	-0.99	8.21	0.30 (AR)
	(0.19)		(0.27)		(1.22)		

Notes: See Table 1.

Table 3. Government investment multipliers with state-dependency based on Auerbach and Gorodnichenko's (2012) continuous recession indicator

Horizon (h)	State-dependent model				*p*-Value for the hypothesis test†
	Expansion		Recession		
	Multiplier	F-statistic	Multiplier	F-statistic	
2	0.45	254.99	0.70	150.61	0.78 (HAC)
	(0.52)		(0.68)		
4	-0.28	192.96	0.06	49.93	0.41 (HAC)
	(0.48)		(0.55)		
6	-0.38	278.34	-0.19	39.56	0.49 (HAC)
	(0.41)		(0.40)		
8	-0.45	218.96	-0.38	43.20	0.69 (HAC)
	(0.30)		(0.35)		
10	-0.49	172.98	-0.55	50.27	0.99 (HAC)
	(0.32)		(0.38)		
12	-0.40	144.72	-0.61	55.55	0.64 (HAC)
	(0.35)		(0.36)		

Notes: †The null hypothesis is that the multipliers are the same across states of expansion and recession. See Table 1.

Table 4. Government consumption multipliers with state-dependency based on Auerbach and Gorodnichenko's (2012) continuous recession indicator

Horizon (h)	State-dependent model				*p*-Value for the hypothesis test†
	Expansion		Recession		
	Multiplier	F-statistic	Multiplier	F-statistic	
2	0.39	302.07	0.37	213.87	0.69 (HAC)
	(0.34)		(0.59)		
4	0.24	105.35	0.38	114.92	0.93 (HAC)
	(0.33)		(0.41)		
6	0.05	58.93	0.17	120.52	0.60 (HAC)
	(0.27)		(0.27)		
8	0.07	36.36	0.07	73.34	0.49 (HAC)
	(0.24)		(0.28)		
10	0.09	26.39	0.01	43.95	0.46 (HAC)
	(0.22)		(0.28)		
12	-0.01	18.74	-0.04	24.30	0.68 (HAC)
	(0.19)		(0.28)		

Notes: See Tables 3 and 1.

Table 5. Government investment multipliers with state-dependency on private debt

Horizon (h)	State-dependent model				*p*-Value for the hypothesis test†
	Low private debt		High private debt		
	Multiplier	F-statistic	Multiplier	F-statistic	
2	0.73	630.79	1.27	253.63	0.50 (HAC)
	(0.60)		(0.41)		
4	0.44	345.32	1.05	137.49	0.39 (HAC)
	(0.40)		(0.49)		
6	0.57	179.90	0.83	69.09	0.70 (HAC)
	(0.35)		(0.49)		
8	0.66	121.43	0.26	50.94	0.36 (HAC)
	(0.30)		(0.31)		
10	0.52	84.08	-0.03	37.11	0.20 (HAC)
	(0.25)		(0.37)		
12	0.37	63.59	-0.02	31.03	0.31 (HAC)
	(0.19)		(0.36)		

Notes: †The null hypothesis is that the multipliers are the same across states of low and high private debt. See Table 1.

Table 6. Government consumption multipliers with state-dependency on private debt

Horizon (h)	State-dependent model				*p*-Value for the hypothesis test†
	Low private debt		High private debt		
	Multiplier	F-statistic	Multiplier	F-statistic	
2	0.26	293.61	0.33	108.33	0.93 (HAC)
	(0.69)		(0.51)		
4	-0.06	90.84	0.33	83.61	0.63 (HAC)
	(0.72)		(0.42)		
6	-0.11	36.09	0.43	65.48	0.41 (HAC)
	(0.59)		(0.36)		
8	0.04	17.79	0.70	43.48	0.28 (HAC)
	(0.56)		(0.31)		
10	0.04	12.95	0.61	35.70	0.33 (HAC)
	(0.52)		(0.28)		
12	-0.15	7.98	0.44	28.92	0.22 (AR)
	(0.45)		(0.21)		

Notes: See Tables 5 and 1.

Table 7. Testing the null hypothesis that the government investment multiplier is the same as the government consumption multiplier at each horizon

Horizon (h)	Linear model Full sample	Linear model Pre-COVID period
1	0.51	0.32
2	0.32	0.25
3	0.24	0.25
4	0.32	0.35
5	0.49	0.38
6	0.49	0.40
7	0.47	0.35
8	0.25	0.28
9	0.20	0.23
10	0.16	0.18
11	0.20	0.21
12	0.26	0.29

Notes: The table reports *p*-values for a Wald test in the linear model. The alternative hypothesis is that the two multipliers are not the same.

Table 8. Government investment multipliers dependent on the unemployment threshold: pre-COVID period 1947Q1 – 2019Q4

Horizon (h)	Linear model (no state-dependency)		State-dependent model				
			Low unemployment		High unemployment		
	Multiplier	F-statistic	Multiplier	F-statistic	Multiplier	F-statistic	*p*-Value for the hypothesis test†
2	0.86**	405.29	1.22	309.24	1.12	81.25	0.92 (HAC)
	(0.38)		(0.44)		(0.97)		
4	0.49*	281.02	1.12	318.24	1.14	31.23	0.98 (HAC)
	(0.25)		(0.41)		(0.91)		
6	0.41**	258.72	1.11	224.97	1.56	16.88	0.62 (HAC)
	(0.19)		(0.46)		(0.75)		
8	0.31	215.67	0.82	134.72	1.53	11.86	0.39 (HAC)
	(0.19)		(0.46)		(0.63)		
10	0.18	168.41	0.53	90.04	1.05	10.08	0.49 (HAC)
	(0.20)		(0.39)		(0.62)		
12	0.06	139.42	0.32	85.38	0.88	8.94	0.54 (AR)
	(0.19)		(0.33)		(0.65)		

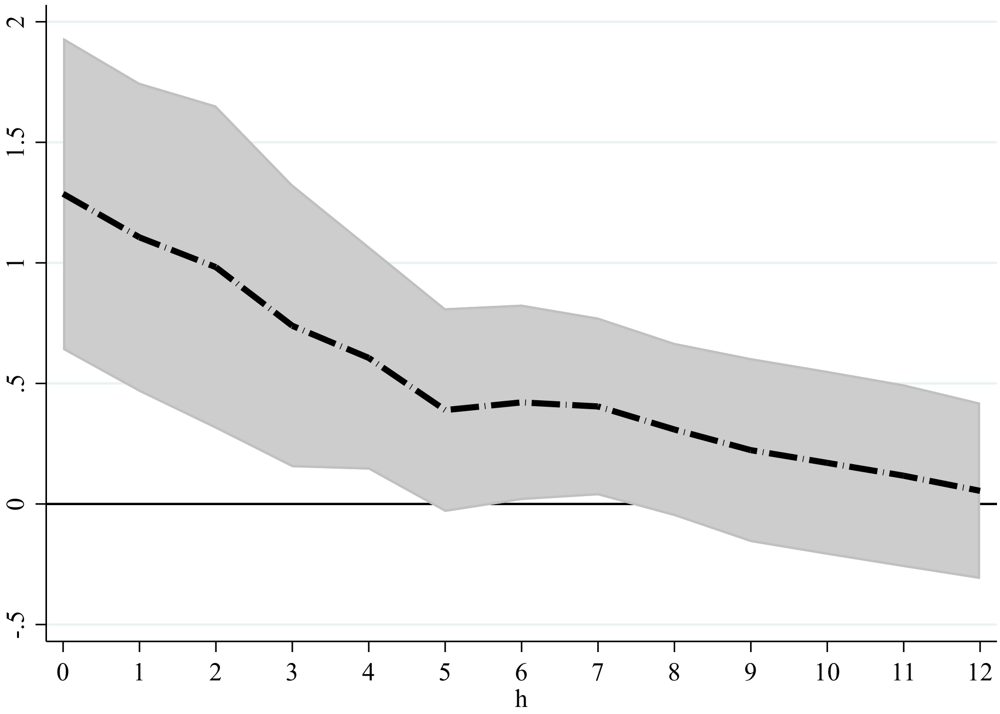
Notes: See Table 1.

Table 9. Government consumption multipliers dependent on the unemployment threshold: pre-COVID period 1947Q1 – 2019Q4

Horizon (h)	Linear model (no state-dependency)		State-dependent model				
			Low unemployment		High unemployment		
	Multiplier	F-statistic	Multiplier	F-statistic	Multiplier	F-statistic	*p*-Value for the hypothesis test†
2	0.36	183.33	0.48	329.34	0.32	76.24	0.84 (HAC)
	(0.34)		(0.40)		(0.80)		
4	0.18	95.41	0.42	112.91	-0.15	24.48	0.31 (HAC)
	(0.30)		(0.33)		(0.53)		
6	0.06	63.15	0.27	48.46	-0.32	26.01	0.34 (HAC)
	(0.26)		(0.29)		(0.53)		
8	0.07	45.70	0.24	31.38	-0.61	28.07	0.30 (HAC)
	(0.24)		(0.30)		(0.69)		
10	0.10	37.40	0.21	26.08	-0.64	18.47	0.41 (HAC)
	(0.21)		(0.29)		(0.88)		
12	0.001	27.16	0.07	18.50	-0.97	8.37	0.31 (AR)
	(0.19)		(0.27)		(1.20)		

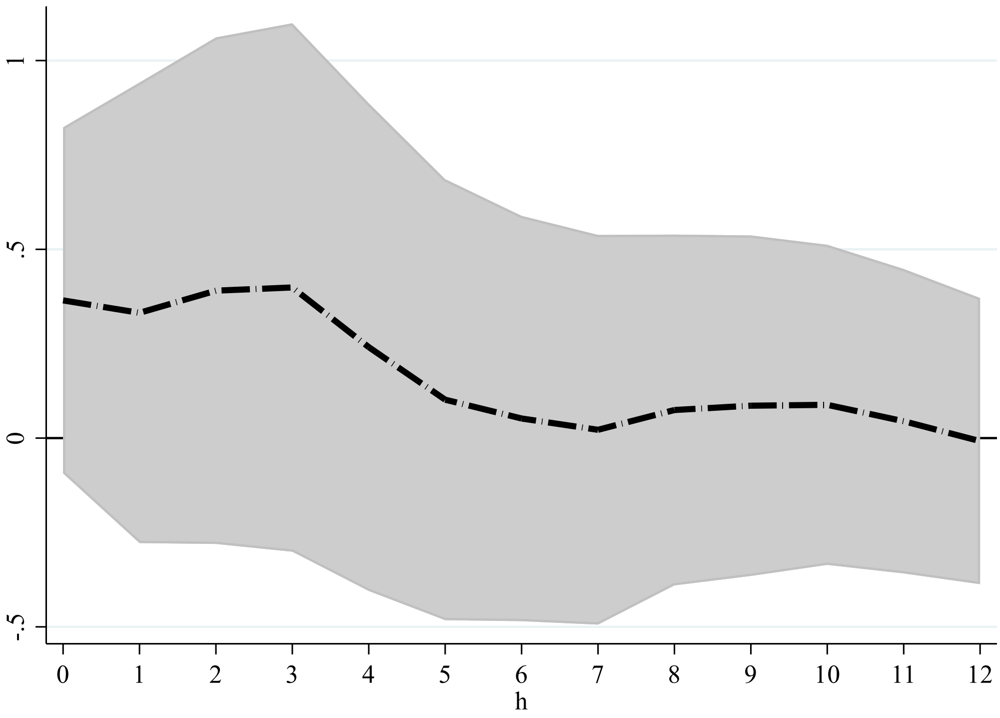
Notes: See Table 1.

Figure 1. Government investment multipliers: linear model



Notes: This figure presents the cumulative impulse responses of real GDP for the linear model, where "h" refers to the horizon. The shaded area shows the 95% confidence intervals based on Newey-West HAC standard errors.

Figure 2. Government consumption multipliers: linear model



Notes: See Figure 1.

Appendix A: Code, Raw Data and Data Transformations

All data are in billions of U.S. dollars or in percent, quarterly, and seasonally adjusted, unless otherwise note. The time period covered is 1947Q1 to 2022Q2. The raw data are mostly from the Federal Reserve Bank of St. Louis' data archive FRED, except for a few different sources listed below. Data were collected on various days in September 2022. The Stata codes and the data are available from the authors on request. The Stata programmes are adapted by the authors from the Stata code made available by Ramey and Zubairy (2018), Supplementary Material, Data Archive, at https://www.journals.uchicago.edu/doi/suppl/10.1086/696277 (last accessed on 12 July 2022), codes *jordagk.do* and *jordagk_ar.do*.

The data definitions and sources follow. The series identifiers refer to the FRED data base if no other source is given. The untransformed raw U.S. economy data employed in this paper are:

- Nominal GDP, series "GDP";
- Nominal government consumption expenditures and gross government investment, series "GCE";
- Nominal gross government investment, series "A782RC1Q027SBEA";
- Chain-type price index for GDP, 2012=100, series "GDPCTPI";
- GDP implicit price deflator, 2012=100, series "GDPDEF";
- Consumer price index (CPI), 1982-1984=100, all urban consumers, monthly, series "CPIAUCSL";
- Nominal government current tax receipts, series "W054RC1Q027SBEA";
- Nominal government current transfer payments, series "A084RC1Q027SBEA;
- Nominal federal government debt, series "FGSDODNS", 1951Q4 to 2022Q2;
- Nominal state and local government debt, series "SLGSDODNS", 1951Q4 to 2022Q2;
- Public debt to GDP ratio, 1947Q1 to 1951Q3, sourced from Bernardini and Peersman (2018), corresponds to (FGSDODNS + SLGSDODNS)/GDP, data are from the data archive, available at http://qed.econ.queensu.ca/jae/datasets/bernardini001/;
- Nominal domestic non-financial-sector debt, securities and loans, liabilities, series "TODNS", 1951Q4 to 2022Q2;
- Private debt to GDP ratio, domestic, non-financial sector, 1947Q1 to 1954Q3, sourced from Bernardini and Peersman (2018), ibid;
- Civilian unemployment rate, percent, monthly, series "UNRATE", 1948:01 to 2022:06; 1947Q1 to Q4 sourced from Ramey and Zubairy (2018), Supplemental Material, at the web address above;

- Nominal federal funds shadow rate, January 1990 to February 2022, monthly, not seasonally adjusted, based on the model of Wu and Xia (2016), sourced from the Federal Reserve Bank of Atlanta, https://www.atlantafed.org/cqer/research/wu-xia-shadow-federal-funds-rate;
- Nominal effective federal funds rate, monthly, not seasonally adjusted, July 1954 to June 2022, series "FEDFUNDS"; 1947Q1 to 1954Q2, quarterly, not seasonally adjusted, sourced from Bernardini and Peersman (2018), ibid;
- COVID-19 stringency index for the U.S. The stringency index is a composite measure based on nine response indicators. It includes data on school closures, workplace closures, and travel bans, rescaled to a value from 0 to 1 (1 = strictest). Daily observations from January 2020 to June 2022 are converted to quarterly averages, sourced from Our World in Data, https://ourworldindata.org/.

All government spending and tax data are for the total government sector as defined in the U.S. Bureau of Economic Analysis' (BEA) National Income and Product Accounts (NIPA): https://apps.bea.gov/iTable/?ReqID=19&step=4&isuri=1&1921=flatfiles#eyJhcHBpZCI6MTksInN0ZXBzIjpbMSwyLDMsM10sImRhdGEiOltbIkNhdGVnb3JpZXMiLCJTdXJ2ZXkiXSxbIk5JUEFfVGFibGVfTGlzdCIsIjg2Il0sWyJGaXJzdF9ZZWFyIiwiMTk0NyJdLFsiTGFzdF9ZZWFyIiwiMjAyMiJdLFsiU2NhbGUiLCItOSJdLFsiU2VyaWVzIiwiUSJdXX0. Hence, federal, state and local government levels are included, net of transfer payments between these entities themselves.

Variables labelled "nominal" above are transformed, as needed, to real variables with the chain-type price index for GDP. Real government consumption purchases are calculated as (GCE – nominal gross government investment)/chain-type price index for GDP. Inflation is based on the quarterly change of the natural logarithm of the CPI, multiplied by 400 to arrive at annual rates. Quarterly CPI data are the average of the monthly observations in each quarter.[12] For the robustness analysis, an alternative measure of the inflation rate is calculated with the GDP-deflator, instead of the CPI, in the same manner. Real government current net tax receipts are defined as (nominal government current tax receipts – nominal government current transfer payments)/chain-type price index for GDP. The government debt to GDP ratio is defined as the sum of nominal federal plus nominal state and local government debt divided by nominal GDP. Private debt is defined as in Bernardini and Peersman (2018) as the ratio of nominal domestic non-financial-sector private debt divided by nominal GDP. The interest rate is represented by the effective federal fund rate for the period before January 1990 and after February 2022, when the zero lower bound of interest rates was not relevant (see https://www.atlantafed.org/cqer/research/wu-xia-shadow-federal-funds-rate).

[12] In the same way, quarterly data are calculated for the unemployment rate and interest rate.

Appendix B: Sensitivity Analysis

Table B.1. Total government spending multipliers dependent on the unemployment threshold

Horizon (h)	Linear model (no state-dependency)		State-dependent model				
			Low unemployment		High unemployment		
	Multiplier	F-statistic	Multiplier	F-statistic	Multiplier	F-statistic	*p*-Value for the hypothesis test†
2	0.48**	169.79	0.51	359.21	0.38	98.02	0.68 (HAC)
	(0.17)		(0.24)		(0.36)		
4	0.29*	115.02	0.48	121.24	0.10	57.70	0.38 (HAC)
	(0.16)		(0.17)		(0.42)		
6	0.14	92.51	0.38	62.67	0.33	44.09	0.90 (HAC)
	(0.13)		(0.15)		(0.42)		
8	0.13	62.47	0.33	47.15	0.27	24.95	0.91 (HAC)
	(0.11)		(0.14)		(0.44)		
10	0.10	50.18	0.23	39.45	0.11	13.07	0.82 (HAC)
	(0.10)		(0.14)		(0.48)		
12	0.02	37.68	0.11	28.81	-0.04	7.58	0.79 (AR)
	(0.10)		(0.14)		(0.56)		

Notes: See Table 1.

Table B.2. Total government spending multipliers with state-dependency based on Auerbach and Gorodnichenko's (2012) continuous recession indicator

Horizon (h)	State-dependent model				
	Expansion		Recession		*p*-Value for the hypothesis test†
	Multiplier	F-statistic	Multiplier	F-statistic	
2	0.43	457.81	0.66	215.76	0.64 (HAC)
	(0.27)		(0.42)		
4	0.15	152.69	0.28	91.43	0.76 (HAC)
	(0.25)		(0.39)		
6	0.09	63.49	0.06	67.36	0.84 (HAC)
	(0.21)		(0.26)		
8	0.06	39.46	-0.06	51.88	0.61 (HAC)
	(0.18)		(0.22)		
10	0.05	28.29	-0.17	50.02	0.44 (HAC)
	(0.17)		(0.22)		
12	-0.01	20.05	-0.24	44.66	0.44 (HAC)
	(0.19)		(0.25)		

Notes: See Table 3.

Table B.3. Total government spending multipliers with state-dependency on private debt

Horizon (h)	State-dependent model				*p*-Value for the hypothesis test†
	Low private debt		High private debt		
	Multiplier	F-statistic	Multiplier	F-statistic	
2	0.36	314.63	0.66	132.11	0.45 (HAC)
	(0.29)		(0.34)		
4	0.09	156.62	0.59	72.84	0.22 (HAC)
	(0.28)		(0.34)		
6	0.11	62.03	0.59	41.62	0.19 (HAC)
	(0.20)		(0.34)		
8	0.22	39.00	0.62	29.35	0.26 (HAC)
	(0.18)		(0.32)		
10	0.21	33.92	0.45	22.58	0.47 (HAC)
	(0.16)		(0.30)		
12	0.10	28.87	0.32	17.61	0.46 (HAC)
	(0.13)		(0.26)		

Notes: See Table 5.

Table B.4. Total government spending multipliers dependent on the unemployment threshold: alternative specifications for the linear baseline model

Horizon (h)	Polynomial trend for the Gordon-Krenn transformation		Adding Ramey and Zubairy's (2018) military news variable to the controls	
	Multiplier	F-statistic	Multiplier	F-statistic
2	0.60**	195.10	0.49**	167.57
	(0.23)		(0.17)	
4	0.42*	146.02	0.29*	109.96
	(0.25)		(0.16)	
6	0.32	119.08	0.15	87.09
	(0.26)		(0.13)	
8	0.33	77.49	0.13	59.50
	(0.26)		(0.11)	
10	0.33	53.65	0.11	48.21
	(0.28)		(0.10)	
12	0.27	24.03	0.03	36.09
	(0.30)		(0.10)	

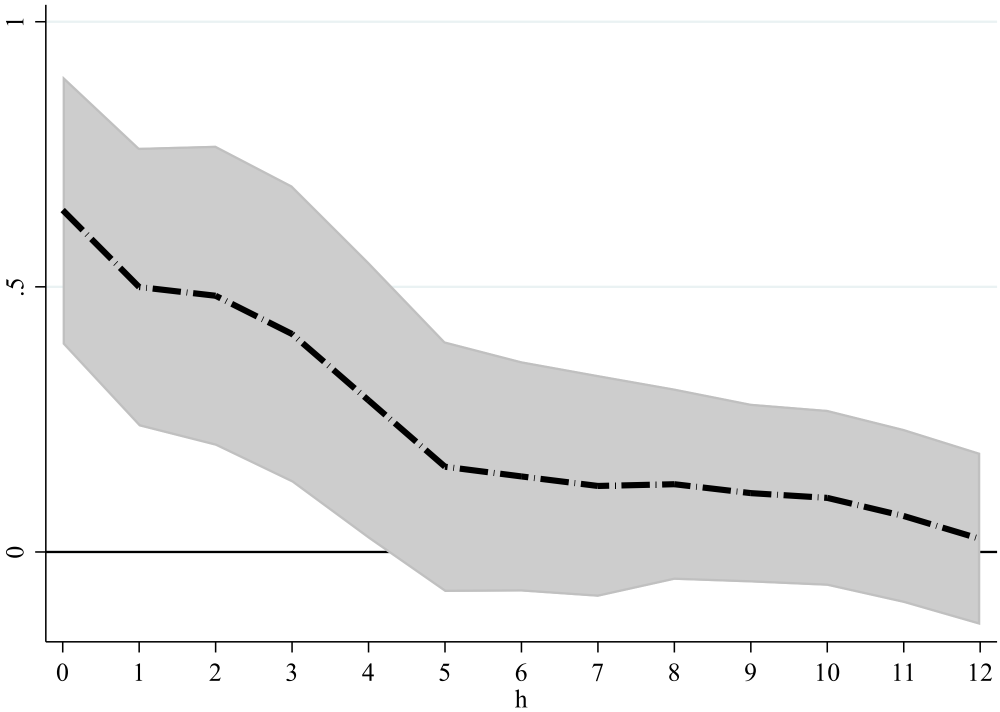
Notes: See Table 1.

Table B.5. Tests and total government spending multipliers dependent on the unemployment threshold with 12 lags on the controls (instead of 4 lags)

Horizon (h)	Linear model (no state-dependency)		*P*-values for tests of the null hypothesis of	
	Multiplier	F-statistic	Equivalence of total government spending multipliers across states	Equivalence of the government investment and government consumption multipliers
2	0.45**	194.94	0.19 (HAC)	0.16 (HAC)
	(0.17)			
4	0.23*	85.46	0.16 (HAC)	0.23 (HAC)
	(0.18)			
6	0.09	56.07	0.51 (AR)	0.30 (HAC)
	(0.16)			
8	0.11	41.20	0.84 (AR)	0.15 (HAC)
	(0.12)			
10	0.06	32.75	0.92 (AR)	0.17 (HAC)
	(0.10)			
12	0.00	26.77	0.87 (AR)	0.34 (HAC)
	(0.09)			

Notes: See Table 1.

Figure B.1. Total government spending multipliers: linear model



Notes: See Figure 1.