Fiscal Multipliers in Advanced and Developing Countries: Evidence from Military Spending

Viacheslav Sheremirov[a,*], Sandra Spirovska[b]

[a] *Federal Reserve Bank of Boston, Research Department T-9, 600 Atlantic Ave, Boston, MA 02210, United States*
[b] *University of Wisconsin, Department of Economics, 1180 Observatory Drive, Madison, WI 53706, United States*

Abstract

This paper provides new evidence on the effects of government spending on output from a large panel of advanced and developing countries. We identify government spending shocks using variation in international military spending. We estimate a one-year fiscal multiplier, pooled across all countries in the sample, in the vicinity of 0.8. The pooled cumulative multiplier reaches its peak at 0.86 two years after the shock. We find substantial heterogeneity across economic environments as well as across countries: The multipliers are relatively large (above one) in advanced economies, in recessions, for negative shocks, under a fixed exchange rate, and in closed economies. We also analyze scenarios in which the identifying restrictions may not hold, highlighting the possible limitations of the military-spending approach and providing a wider range of possible effects.

Keywords: fiscal policy, military spending, multiplier

JEL: E62, F44, H56, H87, O23

[*] This draft: July 2021.

Email addresses: viacheslav.sheremirov@bos.frb.org (Viacheslav Sheremirov), spirovska@wisc.edu (Sandra Spirovska)

1. Introduction

How does government spending affect output in different economic environments? While much evidence on this question comes from the United States and other developed countries, less is known about the fiscal multiplier in developing countries.[1] Many developing countries are prone to extreme economic fluctuations and deep crises with a high degree of economic slack, and fiscal policy may be relatively more effective in such environments (e.g., Parker, 2011; Michaillat, 2014). Further, to anchor inflation expectations and to reduce uncertainty in the tradable sector, developing countries often choose to peg their currency's value to another currency or a basket of currencies. Fiscal policy operates differently under a peg than under a float, as many theoretical models suggest. Developing countries also differ from advanced economies in the degree of openness to foreign trade and investment, independence of their central banks, reliance on commodities, the quality and type of governance, and the amount of fiscal space available to finance public spending. Thus, it is not a priori clear whether fiscal policy is more or less effective in developing countries and why.

In this paper, we provide new evidence on the output responses to government spending in advanced and developing countries. Using multiple sources, we compile data on output, government spending, military spending, unemployment rates, trade shares, and many other variables for 129 countries during the period 1988–2013. These data enable us to differentiate between the multipliers not only in advanced and developing countries but also in different states of the economy and the business cycle, under different exchange-rate regimes, and for different degrees of openness to trade.

To estimate fiscal multipliers, we combine a local projections method with instrumental variables (IV), using military spending shocks as an instrument for changes in government spending. Thus, our identification strategy requires exogenous variation in military spending (e.g., due to geopolitical and other factors not related directly to economic activity) that affects output only through government spending. This approach has long been used for the United States (e.g., Barro,

[1] Notable studies focusing on the United States include, among others, Blanchard and Perotti (2002), Barro and Redlick (2011), Auerbach and Gorodnichenko (2012), and Ramey and Zubairy (2018). Evidence from other developed countries is documented by Alesina and Ardagna (1998), Perotti (2008), Leigh et al. (2010), Beetsma and Giuliodori (2011), Auerbach and Gorodnichenko (2013), and others.

1981; Hall, 1986; Rotemberg and Woodford, 1992) but not for many other countries. Thus, we analyze the predictors of government spending as well as military spending in our sample, including narrative evidence from specialized reports on military spending. Because the identifying assumptions may be violated for some episodes of military spending, we conduct numerous robustness exercises, providing a range of possible effects.

Our paper makes four main contributions. First, we provide a detailed analysis of military spending in international data, focusing on the scope and limitations of using military spending as a potential instrument for government spending. To the best of our knowledge, this study is the most detailed analysis, to date, of international military spending with a focus on fiscal multipliers in a very large sample of countries. Second, this study provides a unique comparison of advanced economies with a *broad* sample of developing countries in terms of the effects of fiscal policy. Our sample contains a large number of low-income, middle-income, and upper-middle-income countries. This is in contrast to previous studies, which either compare advanced economies with mostly upper-middle-income countries (e.g., Ilzetzki et al., 2013) or focus on a broad sample of developing countries using an identification method that cannot be extended to advanced economies for direct comparison (e.g., international aid disbursements; see Kraay, 2012, 2014). Our study combines the cross-group comparability of the former approach with the developing-sample breadth of the latter. Third, this paper provides novel evidence on the state-dependent and asymmetric effects of fiscal shocks, extending to international data the recent influential literature on differential effects of positive and negative spending shocks (e.g., Barnichon et al., 2021). Fourth, this paper enhances our understanding of open-economy effects of fiscal policy. Not only do we provide novel evidence on exchange-rate regimes and openness to trade from international military spending, but we examine the role of commodity imports and exports in the transmission of fiscal shocks.

We document several empirical results. When advanced and developing countries are combined, the one-year multiplier is estimated at around 0.8. For two years after the shock, the cumulative multiplier remains significantly different from zero but not from one.[2] When the two country

[2]These are two common benchmarks considered in the literature to discriminate across theoretical models. For example, in neoclassical models with a high elasticity of intertemporal substitution and a steep marginal-cost curve, the wealth effect offsets most of the stimulus, resulting in multipliers close to zero (Aiyagari et al., 1992). In New Keynesian models, on the contrary, sticky prices and wages allow for multipliers close to one, as long as monetary policy is accommodative (Woodford, 2011). To obtain multipliers larger than one, however, such models require

groups are considered separately, the baseline multiplier in advanced economies is above 1.5 and about twice as large as it is in developing countries. Due to pronounced heterogeneities within each group, however, we cannot reject the equality of the two multipliers at conventional significance levels. The difference is, nonetheless, economically relevant. Our further analysis indicates that the difference between advanced and developing countries is likely due to small multipliers in least-developed countries, as high-income and upper-middle-income countries have multipliers of comparable magnitudes.

To understand the factors behind heterogeneity in the effects of government spending, we examine the state of the economy and the business cycle. Our major finding is that government spending is more effective in recessions than in expansions. The one-year response of output to a unit shock in government spending is 1.7 in recessions and 0.3 in expansions, with statistically significant differences at horizons of one and two years. We also find larger multipliers during episodes of slack, when the unemployment rate is relatively high, than when unemployment is low. While, similar to the recent literature, we find that negative shocks have a larger effect on output than do positive shocks, the large multipliers in recessions cannot be explained only by negative shocks taking place in recessions and positive shocks in expansions. The incidence of positive and negative shocks is comparable across the regimes, and, most importantly, we find that recessions are associated with larger multipliers even when we focus on negative shocks only. These results suggest that procyclical fiscal policy documented in many developing countries (e.g., Gavin and Perotti, 1997; Talvi and Végh, 2005) may be an important source of recession amplification.

Next, we consider open-economy aspects of fiscal policy. We find that countries that peg the value of their currency have on average larger multipliers than countries that let the exchange rate float. We also find that countries with small shares of gross international trade in output—referred to as closed economies—have a larger multiplier than do open economies. The difference between the multipliers in open and closed economies is statistically significant on impact, whereas the difference between the multipliers for pegs and floats, while quantitatively large, is not statistically significant, due to relatively wide confidence bands. When we focus on open economies only, the difference between the multipliers for pegs and floats becomes quantitatively larger and the

additional frictions (e.g., Galí et al., 2007; Christiano et al., 2011).

significance level gets closer to the 10 percent threshold. While we observe variation in exchange-rate regimes across open economies, closed economies in our sample typically choose not to peg their currency, which may hinder our ability to distinguish between the pegs and the floats in the full sample. The quantitative differences between the multipliers based on an exchange-rate regime and openness to trade are consistent with textbook open-economy models. We conclude that, overall, the effectiveness of fiscal policy depends crucially on the economic environment in place and related policies. Therefore, multiplier estimates obtained from a large panel of countries, a long period, or both are likely less informative about the success or failure of a particular stimulus program than multipliers obtained from homogenous samples.

To address concerns about our identification strategy, we conduct a battery of tests wherein we exclude the cases with potential violation of the exclusion restrictions. To this end, we examine financial crises and wars, commodity prices and exports, anticipation of military spending, and countries with large military imports. We also examine cases of potential instrument weakness due to the small size of a country's military sector. In addition, we consider a small subset of countries for which the exogeneity of military spending may be justified by the significant geopolitical role that they play in international relations or by their commitments as members of an international military alliance. The estimates obtained from these alternative samples are typically in line with our baseline.

Our paper is related to the literature studying the effects of fiscal policy in a large panel of countries. Relative to Ilzetzki et al. (2013), we have a broad coverage of lower-middle-income and low-income countries. We also compare multipliers based on a recursive vector autoregression (VAR) identification with those based on the military-spending IV, and find some important differences even when we account for the sample composition. Relative to Kraay (2012, 2014), we use an identification strategy that can be extended to advanced economies, and therefore we can compare responses in the two groups of countries using the same method. Our data and method are similar to those in Miyamoto et al. (2019), but they study consumption responses, with a focus on the exchange-rate transmission mechanism, whereas this paper focuses on the output multipliers. Because international investment data lack in coverage, we cannot simply infer output responses from consumption responses and therefore need to employ output data directly. Our paper is also related to the literature that relies on cross-sectional variation in government spending (e.g., Naka-

mura and Steinsson, 2014; Suárez Serrato and Wingender, 2014, 2016; Shoag, 2016; Dupor and Guerrero, 2017; Chodorow-Reich, 2018). However, we focus on variation across countries rather than across regions and estimate conventional fiscal multipliers rather than local multipliers.

The paper proceeds as follows. Section 2 describes the data and their sources. Section 3 presents our empirical specification. Section 4 discusses in detail our identification strategy, examining military spending as an instrument for total government spending in international data. Section 5 provides estimates of cumulative multipliers, including those for advanced and developing countries, in recessions and expansions, in closed and open economies, and under different exchange-rate regimes. In Section 6, we summarize numerous robustness checks and tests of our identifying assumptions. Section 7 concludes.

2. Data

To conduct our analysis, we compile data on real GDP, government spending, military spending and other relevant variables for 129 countries during the period 1988–2013. Countries with fewer than half of the observations available for this period are not included in the dataset. This approach balances the time-series and cross-sectional dimensions by focusing on countries with a relatively long history. It also provides a systematic way to exclude several war-torn countries, such as Afghanistan and Iraq, for which many data points are missing. The scope is limited to annual observations.

We obtain data from multiple sources. Spending on current military forces and activities comes from the *Military Expenditure Database*, provided by Stockholm International Peace Research Institute (SIPRI). For a subset of countries, total military spending is decomposed into spending on durables and nondurables/services. The data on real GDP and government spending are taken from the *National Accounts Main Aggregates Database*, compiled by the United Nations' Statistics Division. Government spending is measured by general-government final consumption expenditure.[3] Each series is in constant local-currency units. The unemployment and trade data come from the World Bank's *World Development Indicators*. The baseline development classification is from the International Monetary Fund. Accordingly, 36 countries in our sample are classified

[3]Government investment data are not readily available for a large panel of countries.

Table 1: The Relative Size and Volatility of Government Spending

	$\frac{g}{y}$, %	$\frac{g^m}{g}$, %	$\frac{\sigma(\Delta \log g)}{\sigma(\Delta \log y)}$	$\frac{\sigma(\Delta \log g^m)}{\sigma(\Delta \log g)}$	# Countries
	(1)	(2)	(3)	(4)	(5)
Full sample	16.6	16.3	2.1	3.1	129
	(6.1)	(12.1)	(1.9)	(2.1)	
Advanced economies	19.0	11.9	0.9	3.6	36
	(4.4)	(9.0)	(0.4)	(1.5)	
Developing countries	15.6	18.0	2.5	2.9	93
	(6.4)	(12.8)	(2.0)	(2.3)	

Notes: Columns 1 and 2 show the average shares of government spending (g) in output (y) and of military spending (g^m) in total government spending, respectively. Columns 3 and 4 show the average ratios of standard deviations across time (σ) for the corresponding variables. Standard deviations across countries are in parentheses. Column 5 shows the number of countries.

as advanced and 93 countries as developing. The exchange-rate classification follows Klein and Shambaugh (2008), extended to 2013. To control for countries' military engagement, we rely on data from the Correlates of War Project (Sarkees and Wayman, 2010) and the *UCDP/PRIO Armed Conflict Dataset*, compiled jointly by the Uppsala Conflict Data Program and the Peace Research Institute Oslo (Gleditsch et al., 2002). We supplement this core dataset with additional data on financial crises, governance, arms and commodities trade, oil prices, taxes, and monetary-policy rates. A detailed description of variables used and the corresponding data sources, as well as the list of countries included in the sample and their underlying characteristics, can be found in Online Appendix A.

Table 1 presents some descriptive statistics on the level and variability of government spending and military spending. In the full sample, the average share of government spending in GDP is 16.6 percent (column 1), while the share of military spending in total government spending is 16.3 percent (column 2). The two numbers are similar to the corresponding shares in the United States. The government-spending share in advanced countries is somewhat higher than in developing countries (19.0 percent versus 15.6 percent), while the military-spending share is lower in advanced countries (11.9 percent versus 18.0 percent).

Government spending is more volatile than output in the full sample and in the sample of developing countries but less volatile in the sample of advanced countries (column 3). Military spending is, on average, 3.1 times more volatile than total government spending (column 4). Military spending is more volatile than government spending in both country groups. A lack of variation in postwar U.S. data has often been considered an impediment to the effectiveness of the military-

spending approach (e.g., Hall, 2009). This concern appears less important in international data. Hence, the large share of military spending in developing countries makes it a promising candidate to be a relevant instrument. We return to this discussion in Section 4.1, where we present a formal test for instrument relevance.

3. Empirical Specification

To estimate the cumulative government-spending multiplier, we combine the local projections method (Jordà, 2005) with an IV approach, similar to Ramey and Zubairy (2018). Specifically, we estimate the response of cumulative GDP to cumulative government spending instrumented with the military-spending shock. To capture information available at time t, we control for lags of GDP, government spending, and military spending. We also control for country and time fixed effects. Following Gordon and Krenn (2010) and others, we normalize output and government spending by trend GDP. We estimate the trend as a quadratic time polynomial. The choice of a quadratic polynomial is motivated by the annual frequency and relatively short sample period. Since the local projections method gives rise to autocorrelated errors, we cluster standard errors at the country level.

Specifically, for each horizon $h = 1, 2, \ldots$, we estimate the following specification:

$$\sum_{j=0}^{h-1} y_{i,t+j} = \alpha_{i,h} + \mu_h \sum_{j=0}^{h-1} g_{i,t+j} + \boldsymbol{\psi}_h'(L)\, \mathbf{x}_{i,t-1} + \boldsymbol{\gamma}_h'\, \mathbf{z}_{i,t} + \delta_{t,h} + \varepsilon_{i,t+h-1,h}, \tag{1}$$

where $y_{i,t}$, and $g_{i,t}$ are normalized real GDP and total government spending, respectively, in country i and year t. We instrument $\sum_{j=0}^{h-1} g_{i,t+j}$ with military-spending shocks, $g^{\mathrm{m}}_{i,t}$. Vector $\mathbf{x}_{i,t-1} \equiv (y_{i,t-1},\, g_{i,t-1},\, g^{\mathrm{m}}_{i,t-1})'$ controls for information available at time t, while vector $\mathbf{z}_{i,t}$ contains contemporaneous control variables (e.g., a war indicator). A lag polynomial vector $\boldsymbol{\psi}_h(L)$ collects loadings on $\mathbf{x}_{i,t-1}$. We choose the number of lags using the Akaike and Schwarz information criteria.[4] Vector $\boldsymbol{\gamma}_h$ collects loadings on $\mathbf{z}_{i,t}$. Country and time fixed effects are denoted by $\alpha_{i,h}$ and $\delta_{t,h}$, respectively. The error term $\varepsilon_{i,t+h-1,h}$ corresponds to the second-stage regression estimated at horizon h. The coefficient μ_h in specification (1) measures the cumulative government-spending multiplier at an

[4] In practice, we often need to balance the number of lags with the samples size, especially in uneven subsamples. To obtain a reasonable sample size, in such cases, we use one lag.

h-year horizon.

To measure the multipliers in subsamples (e.g., advanced versus developing countries), and to test the statistical difference between such multipliers, we interact the regression coefficients with a subsample dummy ($\mathbb{I}_i$), estimating the following equation:

$$\begin{aligned}\sum_{j=0}^{h-1} y_{i,t+j} = \mathbb{I}_i \times \left(\alpha_{i,h}^{\mathrm{A}} + \mu_h^{\mathrm{A}} \sum_{j=0}^{h-1} g_{i,t+j} + \boldsymbol{\psi}_h^{\mathrm{A}\prime}(L)\, \mathbf{x}_{i,t-1} + \boldsymbol{\gamma}_h^{\mathrm{A}\prime}\, \mathbf{z}_{i,t} + \delta_{t,h}^{\mathrm{A}} \right) \\ + (1 - \mathbb{I}_i) \times \left(\alpha_{i,h}^{\mathrm{B}} + \mu_h^{\mathrm{B}} \sum_{j=0}^{h-1} g_{i,t+j} + \boldsymbol{\psi}_h^{\mathrm{B}\prime}(L)\, \mathbf{x}_{i,t-1} + \boldsymbol{\gamma}_h^{\mathrm{B}\prime}\, \mathbf{z}_{i,t} + \delta_{t,h}^{\mathrm{B}} \right) + \varepsilon_{i,t+h-1,h},\end{aligned} \tag{2}$$

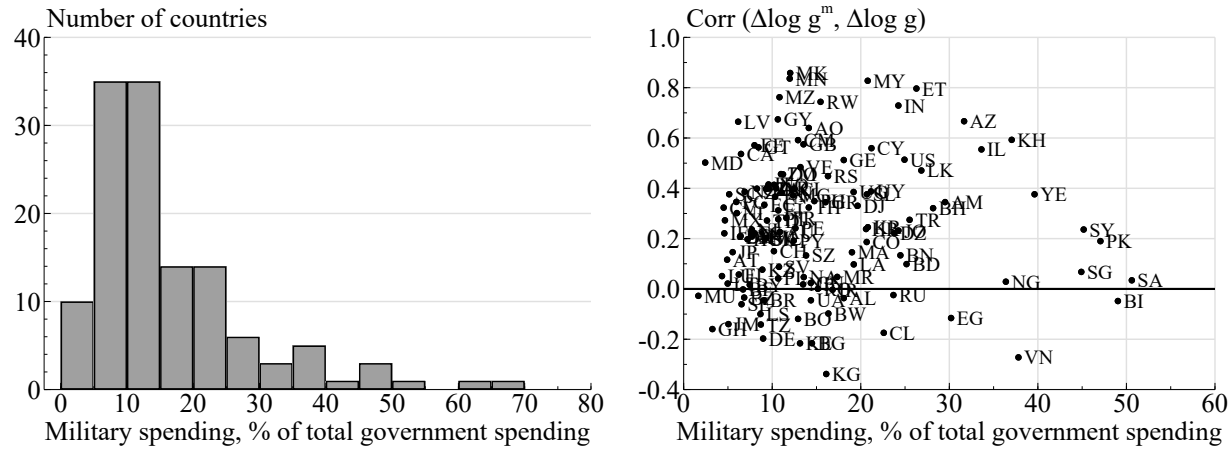
where $\mathbb{I}_i \times \sum_{j=0}^{h-1} g_{i,t+j}$ and $(1-\mathbb{I}_i) \times \sum_{j=0}^{h-1} g_{i,t+j}$ are instrumented with $\mathbb{I}_i \times g_{i,t}^{\mathrm{m}}$ and $(1-\mathbb{I}_i) \times g_{i,t}^{\mathrm{m}}$. In specification (2), μ_h^{A} and μ_h^{B} are the horizon-h cumulative multipliers for subsample A and subsample B, respectively. The test for the hypothesis $\mu_h^{\mathrm{A}} = \mu_h^{\mathrm{B}}$ is straightforward.

4. Identification Strategy

Our identification strategy is based on two conditions: First, military spending does not correlate with the unobserved determinants of output, $\mathrm{cov}(g_{i,t}^{\mathrm{m}}, \varepsilon_{i,t}) = 0$ (*instrument validity*). Second, military spending shocks are not completely offset by changes in other components of government spending, $\mathrm{cov}(g_{i,t}^{\mathrm{m}}, g_{i,t}) > 0$ (*instrument relevance*). In this section, we provide a detailed discussion of these conditions in the context of international data. In Section 4.1, we present supporting evidence as well as a formal statistical test for instrument relevance. In Section 4.2, in the spirit of Granger causality tests, we examine whether military spending correlates with lags of economic variables. In Section 4.3, we provide narrative evidence that, in many cases, large military buildup is due to geopolitical factors. In Section 4.4, we discuss whether our approach would overestimate or underestimate the multiplier for some scenarios of identification failure.

4.1. Instrument Relevance

Figure 1a shows the distribution across countries of the average shares of military spending in total government spending. Only 10 countries have military-spending shares below 5 percent. In more than half of the countries in our sample, military spending makes up more than 10 percent of total government spending. While the shares typically fall within the 5 percent to 25 percent range, the distribution has a long right tail.



(a) The Cross-Country Distribution of Military Spending

(b) Correlation with Government Spending

Figure 1: The Relevance of Military-Spending Shocks across Countries. In the right panel, countries are labeled by their ISO2 codes. For better visibility, we drop four outliers from the right panel: two countries with average $g^{\mathrm{m}}/g > 60$ percent (Oman, United Arab Emirates) and two countries with $\mathrm{Corr}(\Delta \log g^{\mathrm{m}}, \Delta \log g) < -0.4$ (Lebanon, Malta). We verify that these observations do not have a material effect on the slope or the fit of the corresponding linear regression.

Figure 1b shows a scatterplot of these shares (horizontal axis) against the time-correlation between log-changes in military spending and log-changes in total government spending (vertical axis). The correlation is at least moderately positive for a vast majority of countries. Moreover, we often observe a high correlation between military spending and government spending for countries with an average size of military expenditure.

To analyze the relevance condition formally, we employ the Montiel Olea and Pflueger (2013) robust test for weak instruments. This test is robust to heteroskedasticity and the serial correlation in the error term arising from local projections at longer horizons. It is based on the first-stage regression of the cumulative government spending, $\sum_{j=0}^{h-1} g_{i,t+j}$, on the military-spending shock, $g^{\mathrm{m}}_{i,t}$, and the set of control variables from the second stage. While conventional tests require an F-statistic above 10 (e.g., Staiger and Stock, 1997), this test is conservative, with a 5 percent significance threshold above 20. In Section 5.1, we show that our instrument passes the relevance test at the peak horizons in the full sample, and separately in the samples of advanced and developing countries. We report effective F-statistics separately for every subsample analyzed in this paper.

Table 2: Predictors of Government Spending

	Military spending							Total spending		
	Full sample					Adv.	Dev.	All	Adv.	Dev.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
War	0.443***				0.412***	0.144	0.509***	0.307	0.255	0.403
	(0.138)				(0.129)	(0.089)	(0.168)	(0.352)	(0.341)	(0.466)
Output		−0.003			−0.001	0.003	−0.002	−0.062***	−0.096***	−0.055***
		(0.005)			(0.005)	(0.008)	(0.006)	(0.017)	(0.022)	(0.019)
Recession			0.080		0.060	0.119	0.063	0.705***	0.932***	0.647*
			(0.061)		(0.064)	(0.073)	(0.086)	(0.237)	(0.207)	(0.329)
Exchange rate peg				0.001	0.008	−0.084	0.033	0.060	0.632**	−0.102
				(0.068)	(0.061)	(0.161)	(0.064)	(0.205)	(0.300)	(0.244)
Observations	2,899	2,899	2,849	2,806	2,763	846	1,917	2,763	846	1,917

Notes: This table presents estimates from the regression of military spending (columns 1–7) and total government spending (8–10) on the first lag of variables in rows. All specifications control for country and time fixed effects. Standard errors clustered by country are in parentheses. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

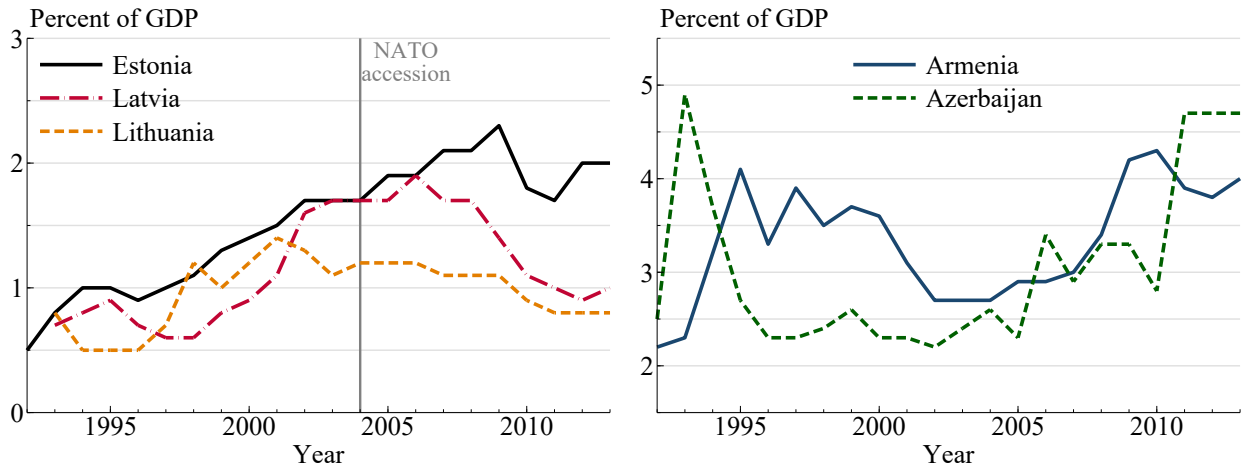
4.2. Predictors of Government Spending

We examine whether military spending can be predicted by the lags of economic as well as geopolitical variables. For this purpose, we regress military spending on the lags of the war indicator, output, recession, and the exchange-rate regime. Columns 1–4 of Table 2 present estimates when we use one predictor at a time, and column 5 when we use all predictors together. The war indicator is the only significant predictor of military spending, and the coefficients on economic variables are small. When we estimate the model separately for advanced and developing countries (columns 6–7), the results are similar, with the only material difference being that the coefficient on wars becomes smaller and insignificant in the advanced countries sample.

Next, we compare these results with a similar model, in which we use total government spending as a left-hand-side variable. In this exercise (columns 8–10), we find that lags of output and recessions have a predictive power both in advanced and developing countries, with the corresponding coefficients being markedly larger than for military spending. The coefficient on the war indicator remains quantitatively large but is not significant at conventional levels.

4.3. Narrative Evidence

Many studies argue that military spending responds to geopolitical events rather than to domestic economic conditions. For example, Collier (2006) and others provide numerous narrative examples of specific episodes in developing countries when military spending reacted to geopolitical events. In the context of our sample, we supplement these studies with additional narrative evidence. Below, we provide two illustrative examples. One example is the NATO enlargement



(a) Accession to NATO

(b) Nagorno-Karabakh Conflict

Figure 2: Examples of Military Spending, Measured by Percent of GDP, due to Geopolitical Factors

in Eastern Europe, which included Latvia, Lithuania, and Estonia. The other example covers the aftermath of a military conflict between Armenia and Azerbaijan over the Nagorno-Karabakh disputed territory. Due to space constraints, we discuss examples from other regions, such as the Middle East and North Africa, in Online Appendix B.[5]

During our sample period, 12 new members joined the NATO alliance. In order to join, the prospective members had to improve the standards of their armed forces in line with those of the current members and thus to spend more on their military. As an illustration, Figure 2a shows the dynamics of military spending as a share of output in the three Baltic countries. Between gaining independence in the early 1990s and their accession to NATO in 2004, these countries' military spending increased from about 0.5 percent of GDP to nearly 2 percent (1.5 percent in Lithuania). We observe a similar pattern in other new members. We note that during this episode the dynamics of military spending were, by and large, determined by long-term security considerations. In Armenia and Azerbaijan, military spending during our sample period was influenced by the Nagorno-Karabakh conflict (Figure 2b). Following the 1994 ceasefire, military spending declined dramatically in both countries. In the late 2000s, political tensions in the region resumed, making both countries increase their military spending (Stålenheim et al., 2008, pp. 185–189).[6]

[5]Online Appendix C.1 describes military spending dynamics in the full sample.

[6]While the report describes the unresolved conflicts in the region as "the major reason for the rapid increase in military spending" (ibid., p. 187), it also mentions economic growth and rising oil revenues as facilitating factors. As described in Section 3, all variables in our analysis are normalized by trend GDP, which accounts for economic

4.4. Violation of the Exclusion Restriction

While in the robustness section we attempt to account for many possibilities when the identification assumptions may fail, ultimately it is not possible to verify these assumptions for each and every country and episode. To put things in perspective, we discuss implications of possible endogeneity of military spending for the magnitude of our estimates. Let $\tilde{g}^{\mathrm{m}}_{i,t}$ be military spending orthogonalized with respect to the controls used in specification (1). A standard result holds that the IV estimator of the multiplier ($\hat{\mu}$) converges in probability to the "actual" multiplier (μ°) plus a component that depends on the covariances of $\tilde{g}^{\mathrm{m}}_{i,t}$, $g_{i,t}$, and the error term $\varepsilon_{i,t}$:

$$\hat{\mu} \xrightarrow{\mathrm{p}} \mu^{\circ} + \frac{\mathrm{cov}(\tilde{g}^{\mathrm{m}}_{i,t},\, \varepsilon_{i,t})}{\mathrm{cov}(\tilde{g}^{\mathrm{m}}_{i,t},\, g_{i,t})}. \tag{3}$$

Under our assumptions, the second term on the right-hand side vanishes, and thus the estimator is consistent, $\mathrm{plim}\,\hat{\mu} = \mu^{\circ}$. But if the exclusion restriction fails, our approach would overestimate the multiplier if $\mathrm{cov}(\tilde{g}^{\mathrm{m}}_{i,t},\, \varepsilon_{i,t}) > 0$ and underestimate it if this covariance is strictly negative, because the relevance condition presumes that $\mathrm{cov}(\tilde{g}^{\mathrm{m}}_{i,t},\, g_{i,t})$ is positive.

We focus on two scenarios of identification failure, which could have the opposite effects on our estimates. One possibility is that military spending, similar to other components of government spending, could be used as a tool for countercyclical fiscal policy. In this case, $\tilde{g}^{\mathrm{m}}_{i,t}$ would increase when the policymaker receives a negative news about current or future economic conditions: that is, $\mathrm{cov}(\tilde{g}^{\mathrm{m}}_{i,t},\, \varepsilon_{i,t}) < 0$. By (3), we would underestimate the government-spending multiplier. This scenario likely applies more to advanced economies with developed policy institutions, as fiscal policy in many developing countries is known to be procyclical (e.g., Talvi and Végh, 2005). Note, however, that this channel works in the direction opposite to our conclusions, as we document relatively large multipliers in advanced countries. This scenario is also more likely to apply to recessions, when countercyclical policies are used to stimulate the economy, than to expansions, when government spending depends on longer-term considerations. This downward bias again works in the direction opposite to our finding of large multipliers in recessions.

Another possibility is that $\mathrm{cov}(\tilde{g}^{\mathrm{m}}_{i,t},\, \varepsilon_{i,t}) > 0$ due to limited fiscal space. This scenario could ap-

growth. We show that our results are not driven by commodity exporters or oil prices in Section 6.4.

ply to developing countries that cannot borrow freely in financial markets. In such a case, military spending would be constrained by contemporaneous tax revenues, which increase in good times and plummet in bad. The positive covariance above can also emerge when, due to rapid economic development, a country strives to play a larger geopolitical role. For example, China got more involved in international affairs during a period of sustained growth. In such cases, our approach may overestimate the effect of government spending on output. Conditional on large multipliers in advanced economies, this source of bias would make the difference between advanced and developing countries even larger than documented in this paper. Hence, both of these possible biases work in the direction opposite to our qualitative conclusions.

5. Results

We now present our main results. Section 5.1 shows estimates obtained in the full sample, and separately for advanced and developing countries. It also provides additional evidence from a small sample of members of a military alliance and from large geopolitical players. Section 5.2 presents estimates of state-dependent multipliers, which differ in recessions and in expansions. It also discusses the asymmetric effects of spending increases and decreases, and studies the interaction between negative shocks and recessions. In Section 5.3, we address open-economy aspects of fiscal policy, focusing on the exchange-rate regimes and openness to trade.

5.1. Baseline Multipliers

5.1.1. Full Sample

We start by presenting our estimates for the full sample. Figure 3 shows the IRFs of government spending and output to a military-spending shock normalized as 1 percent of GDP. The government-spending responses remain positive and significant for four years and decline to zero thereafter. The contemporaneous response is close to one, indicating high passthrough of military spending. Output also responds positively to the military shock, and the response is significant for two years. Thus, the cumulative multiplier combines the persistent increase in output with the persistent increase in government spending. The timing of the output response is, in general, consistent with previous studies (e.g., Barro and Redlick, 2011).

Our estimates of the cumulative multiplier for different combinations of country and time fixed effects are presented in columns 1–4 of Table 3. In all specifications, we control for the war

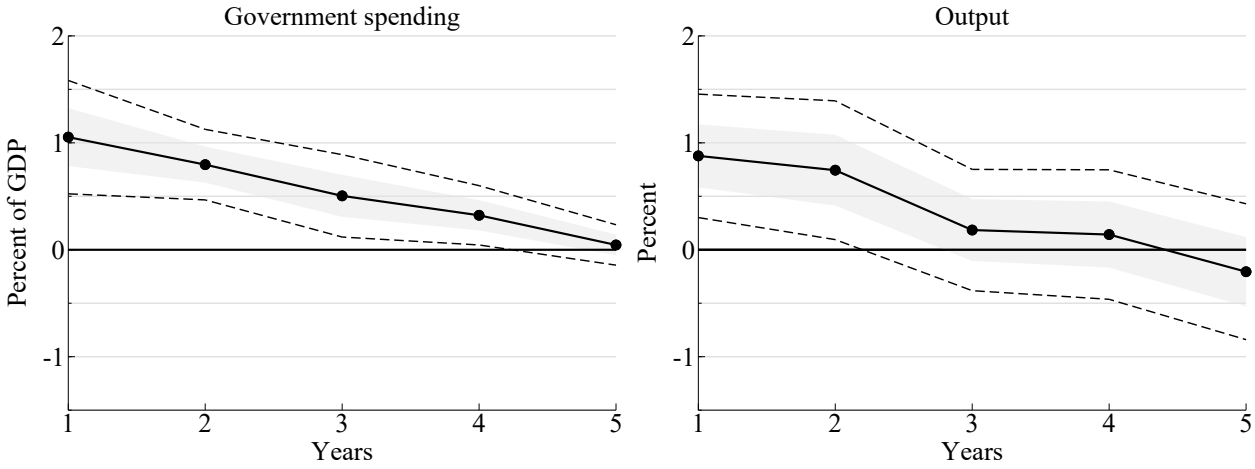


Figure 3: The Responses of Government Spending and Output to a Military-Spending Shock. The figure shows the IRFs of government spending and output to a military spending shock of 1 percent of GDP, estimated using local projections. The gray areas represent one standard error (clustered by country) on each side of the estimates, and the dashed lines show 95 percent confidence bounds.

Table 3: Cumulative Multipliers in Advanced and Developing Countries

	Full sample				Advanced	Developing	p-value
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
One year	0.737**	0.752**	0.799**	0.834**	1.679**	0.814**	
	(0.331)	(0.361)	(0.341)	(0.382)	(0.727)	(0.389)	0.296
	[17.2]	[15.2]	[17.0]	[15.2]	[91.3]	[14.3]	
Two years	0.762**	0.747**	0.823**	0.861**	1.877**	0.818**	
	(0.333)	(0.374)	(0.337)	(0.394)	(0.933)	(0.393)	0.298
	[33.3]	[27.4]	[32.6]	[27.7]	[90.3]	[27.4]	
Three years	0.501*	0.361	0.571*	0.546	1.504*	0.499	
	(0.295)	(0.351)	(0.292)	(0.378)	(0.842)	(0.387)	0.280
	[32.8]	[20.9]	[31.9]	[21.4]	[74.5]	[19.7]	
Four years	0.432	0.224	0.533*	0.507	1.457*	0.451	
	(0.307)	(0.418)	(0.307)	(0.456)	(0.835)	(0.477)	0.297
	[33.3]	[17.9]	[32.3]	[18.6]	[45.9]	[16.1]	
Country effects	N	Y	N	Y	Y	Y	
Time effects	N	N	Y	Y	Y	Y	
Observations	2,563	2,563	2,563	2,563	774	1,789	

Notes: Columns 1–4 present estimates of the cumulative multipliers from specification (1) for different combinations of fixed effects, using the full sample of countries. Columns 5 and 6 show estimates of cumulative multipliers in advanced and developing countries, respectively, from specification (2). Column 7 shows the p-values of the test that the difference between the multipliers in advanced and developing countries is zero. The development groups are based on the IMF classification. In all specifications, we control for the war indicator as well as for four lags of normalized GDP, government spending, and military spending. Standard errors clustered by country are in parentheses. The effective first-stage F-statistics are in brackets. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

indicator as well as four lags of military spending, government spending, and output. Our preferred specification is in column 4. During the first two years after the shock, the multiplier is in the range 0.74–0.86, significantly different from zero but not from one at the 5 percent level, based on standard errors clustered by country. After two years, the multiplier declines and is statistically insignificant at conventional levels if country effects are included.

Table 3 also reports the first-stage effective F-statistics (in brackets). The instrument passes the Montiel Olea and Pflueger (2013) relevance test at the two-year horizon (at the 5 percent level) and at the three-year horizon (at the 10 percent level). For all horizons, the first-stage effective F-statistics are above 10, which is a weaker threshold often referenced in the literature.

We also estimate a range of alternative specifications, which—to save space—we summarize only briefly here.[7] For example, our estimates are not particularly sensitive to controlling for more than one year of lags, with a range of estimates at 0.85–0.90 at the two-year horizon. However, controlling for lags does play a significant role in our estimation method. For instance, not including any lags of military spending results in underestimation of the multiplier, because the shock is contaminated by the multi-year spending programs that are likely to have the most effect on output in the year when they are announced. Not controlling for any lags of output, instead, leads to overestimation of the multiplier, because some of the secular increase in output may be erroneously attributed to the spending shock following a similar trend. Hence, it appears important in practice as well as in theory to remove pretrends in all key variables. Our conclusions are also robust to normalization by lag output instead of trend output. This transformation yields multipliers that are similar to our baseline estimates on impact and are somewhat smaller than the baseline over longer horizons.

We also produce alternative estimates of standard errors. As a baseline, we cluster standard errors at the country level. This approach is conservative, allowing for a serial correlation of arbitrary length rather than h lags, as required by local projections. We confirm that Newey–West standard errors with bandwidth H (maximal h considered) are generally smaller than our baseline. While the baseline approach does not account for the cross-country correlation of errors, we note

[7]The detailed results can be found in Online Appendix C.2. Table C.2 examines alternative lengths of the lag polynomials in control variables. Table C.3 examines pretrends by dropping all lags of each control variable, one at a time. Table C.4 considers an alternative normalization by lag output. Table C.5 examines robustness of standard errors.

that in a large, diverse sample of countries, the spatial correlation matrix is likely quite sparse, especially after controlling for time effects. If many pairwise correlation coefficients are in fact zero, attempting to estimate the sample correlations of residuals for all pairs of countries presents a significant tradeoff between robustness and efficiency, especially when $N \gg T$.[8] When we estimate alternatives that allow for cross-country correlations, such as Driscoll and Kraay (1998) or two-way clustering by country and time, the multiplier's estimates remain significant at the 10 percent level for the first two years. The standard error for the two-year multipliers increases from 0.394 to 0.439 and 0.457, respectively, when using these two methods. As we report our results for subsamples, we verify that they remain significant at least at the 10 percent level under these alternative procedures to compute standard errors.

We compare our baseline estimates to those in the literature. For example, using U.S. military news, Ramey and Zubairy (2018) report state-invariant multipliers, cumulated over a two-year period and a four-year period, in the range 0.66–0.71. Using a structural VAR instead, Auerbach and Gorodnichenko (2012) report a 0.57 cumulative multiplier over a five-year period for the United States, while Ilzetzki et al. (2013) report a long-run multiplier of 0.66 for a large sample of high-income countries. While our pooled multipliers look quite similar, despite using annual data for a different period and a much larger sample of countries, we show next that this similarity masks important heterogeneity across countries and episodes.

5.1.2. Advanced and Developing Countries

Next, we estimate the multipliers separately for developed and developing countries, as classified by the International Monetary Fund.[9] The point estimates are larger for developed countries (column 5 of Table 3) than for developing countries (column 6) across all horizons. An additional \$1 of government spending increases GDP on impact by \$1.68 in developed countries and by \$0.81

[8] We note that a detailed analysis of spatial correlations engendered by fiscal shocks requires a separate study and therefore is beyond the scope of this paper. For example, Bräuning and Sheremirov (2019) study international output spillovers from U.S. monetary shocks employing a variety of spatial techniques.

[9] See Table A.1 and the corresponding description in the online appendix. As an alternative, we also use a World Bank development classification based on gross national income. This classification allows us to break down the sample of developing countries further, into middle-income and low-income countries. The multiplier's magnitude in developing countries is pulled downward by low-income countries. Our estimates for low-income countries are similar to those of Kraay (2012), who studies the effects of World Bank aid disbursements in 29 countries, most of which are low income. This and other additional results on subsamples can be found in Online Appendix C.3.

in developing countries. The multiplier in developed countries is large and statistically different from zero at all considered horizons, while in developing countries it becomes smaller and insignificant after two years. The difference in point estimates between the subsamples, however, is not statistically significant at conventional levels (column 7) due to wide confidence bands.[10]

The smaller multipliers are consistent with tight fiscal space in many developing countries, which often have limited access to domestic debt and therefore rely primarily on taxation and foreign debt—including debt from the IMF, which may come with strings attached and take longer to secure. Hagedorn et al. (2019) show that, in a heterogeneous-agents New Keynesian model with incomplete markets, the fiscal multiplier is considerably above one if spending is financed by debt, but below one if taxes are raised. Our empirical estimates appear consistent with the debt-financed multiplier in developed countries and the tax-financed multiplier in developing countries.

Note that the multipliers for advanced countries are larger than the ones reported in the literature discussed earlier. While the standard errors are too large to reject conventional (lower) values, this discrepancy could be explained by differences in the sample period. A significant portion of our sample is influenced by a deep recession and a tepid recovery across the advanced nations during the Great Financial Crisis, which may result in larger multipliers. Comparing our results with the relatively scarce literature on developing countries, we observe that Ilzetzki et al. (2013) document negative but insignificant multipliers for a sample of mostly upper-middle-income developing countries. We, too, find smaller multipliers for developing countries than for developed countries, but in our sample the multipliers are positive and significant.[11] Our estimates are also numerically larger than, but not statistically different from, the ones in Kraay (2014), who estimates a multiplier of about 0.4 in a sample of developing countries similar to ours, but for a different period. While our results are qualitatively similar to Kraay's, our evidence is based on a different identification strategy and a different period.

[10] In Online Appendix B.1, we focus on a smaller sample of countries in military alliances, such as NATO for advanced economies and Islamic Military Counter Terrorism Coalition for developing countries, and reach qualitatively similar conclusions.

[11] However, when we match the sample considered in Ilzetzki et al. (2013) as closely as possible (Online Appendix C.4), our method produces multipliers for developing countries similar to those in advanced economies. This is consistent with the result, described in footnote 9, that there is less difference between advanced and upper-middle-income countries in our data.

Table 4: State-Dependent Multipliers

	Full sample			Developing countries		
	Recessions	Expansions	p-value	Recessions	Expansions	p-value
	(1)	(2)	(3)	(4)	(5)	(6)
One year	1.744***	0.325		1.569**	0.267	
	(0.667)	(0.223)	0.046	(0.616)	(0.207)	0.049
	[38.9]	[10.0]		[39.8]	[9.2]	
Two years	1.536**	0.348		1.297**	0.274	
	(0.615)	(0.303)	0.089	(0.561)	(0.293)	0.116
	[45.2]	[18.1]		[46.9]	[17.9]	
Three years	0.948*	0.044		0.648	–0.059	
	(0.522)	(0.333)	0.149	(0.503)	(0.334)	0.255
	[31.2]	[14.7]		[28.4]	[13.2]	
Observations	500	2,063		322	1,467	

Notes: Recessions in column 1 are defined at annual frequencies as a decrease in real GDP. Estimates in columns 4–6 are for the sample of developing countries. p-values in columns 3 and 6 are for the differences between the point estimates in the two preceding columns. Standard errors clustered by country are in parentheses. The effective first-stage F-statistics are in brackets. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

5.2. State-Dependent Multipliers

5.2.1. Recessions and Expansions

To compare the multipliers in recessions and expansions, we first define a recession at an annual frequency as a decrease in real GDP relative to the previous year.[12] Because the data are observed at an annual frequency, this definition is likely to miss small recessions. We then estimate (2) using the recession indicator. The results are presented in Table 4. We find that, in the full sample, the fiscal multipliers are larger in recessions (column 1) than in expansions (column 2). In recessions, a \$1 increase in government spending, on impact, leads to a statistically significant increase in output of \$1.74, whereas in expansions the increase is \$0.33 and not statistically significant. The recession multiplier remains statistically significant and above or near one for at least three years after the shock. The difference between the recession and expansion multipliers is statistically significant, at least at the 10 percent level, for two years (column 3).[13]

We also estimate the state-dependent multipliers for developing countries only (columns 4–6). The multiplier, on impact, is 1.57 in recessions and 0.27 in expansions. The difference is statistically significant at the 5 percent level. Our recession multipliers are larger than the ones by Kraay (2014), who reports estimates in the range 0.6–0.8. Besides the differences in the identification

[12] Since recessions are usually defined at a quarterly frequency, we experiment with treating a year following the negative growth in GDP as a recession, too. This approach accounts for the cases when strong recovery at the end of a recessionary year offsets the negative growth at the beginning.

[13] To define the state of the economy, we also use the percentiles of the unemployment rate distribution. See Appendix Table C.8.

Table 5: Positive versus Negative Spending

	All countries			Developing countries			Negative shocks		
	Positive	Negative	p-value	Pos.	Neg.	p-val.	Rec.	Exp.	p-val.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
One year	0.290	1.623***		0.273	1.663**		2.650***	0.867**	
	(0.271)	(0.611)	0.037	(0.259)	(0.668)	0.044	(0.838)	(0.366)	0.020
	[28.4]	[12.5]		[27.3]	[10.7]		[11.3]	[8.5]	
Two years	0.280	1.343**		0.238	1.304**		1.810**	0.726*	
	(0.352)	(0.561)	0.115	(0.321)	(0.597)	0.121	(0.885)	(0.422)	0.177
	[47.7]	[15.5]		[46.9]	[15.4]		[19.4]	[9.0]	
Three years	−0.081	0.877*		−0.144	0.861		0.647	0.526	
	(0.338)	(0.528)	0.128	(0.329)	(0.574)	0.122	(0.929)	(0.562)	0.875
	[27.9]	[14.3]		[28.4]	[12.5]		[23.1]	[6.9]	
Observations	1,077	1,486		779	1,010		344	1,142	

Notes: This table shows the multipliers for positive and negative shocks (columns 1–3), as wells as for their interactions with the development groups (4–6) and the states of the business cycle (7–9). p-values in columns 3, 6, and 9 are for the differences between the point estimates in the two preceding columns. Standard errors clustered by country are in parentheses. The effective first-stage F-statistics are in brackets. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

strategies, our data end in 2013 and therefore contain the aftermath of the Global Financial Crisis.

5.2.2. Positive versus Negative Spending

To understand the state-dependent nature of the multipliers, we estimate asymmetric effects of government spending, focusing separately on spending increases and decreases, as in Barnichon et al. (2021). Similar to that paper, we find larger multipliers associated with negative shocks and validate this result for a large sample of developing countries. We also find that the multipliers in recessions are significantly larger than in expansions when we focus only on negative shocks.

These results are presented in Table 5. We note that our sample of shocks is rather balanced, with 42 percent of observations corresponding to increases in military spending. The multiplier associated with expansionary shocks is below 0.3 and statistically indistinguishable from zero at all considered horizons (column 1), whereas the contractionary multiplier is 1.6 on impact, remaining above one for two years and significantly different from zero (at least at the 10 percent level) for three years (column 2). The expansionary and contractionary multipliers are different from each other on impact at the 5 percent level (column 3). We obtain similar estimates when we partition the sample based on increases and decreases in total government spending, rather than military spending shocks.

The asymmetry of government-spending effects holds in the sample of developing countries. The multiplier associated with spending decreases in developing countries is larger than the multiplier for increases (columns 4–6). In fact, the multiplier for contractionary shocks in developing

countries, at 1.66 on impact, is close to the overall multiplier in advanced economies.[14] We note that, due to using annual data for a relatively short period, the size of the advanced-economies sample is too small to measure the multipliers accurately in the subsamples.[15] Another reason that we focus on a broad sample of developing countries is that there is significantly less evidence in the literature on state-dependence and asymmetric effects of fiscal policy from lower-middle-income and low-income countries, which represent a large part of our sample.

Finally, we estimate the multipliers in recessions and expansions for negative shocks only (columns 7–9). As in the full sample, we find much larger multipliers in recessions than in expansions in this case. Note also that about three-fourths of all negative shocks take place in expansions, a proportion that is similar to the incidence of expansions in our sample (about 80 percent). Hence, the sign of the shocks and the state of the economy provide two different sources of nonlinear effects of government spending.[16]

5.3. Multipliers in the Open Economy

5.3.1. Exchange-Rate Regimes

To evaluate the effect of the exchange-rate regime on the multiplier size, we split the sample into pegs and floats. Under a fixed exchange rate, the impact multiplier is 1.20 and significant at the 10 percent level, compared with 0.45 and insignificant under a flexible rate (columns 1 and 2 of Table 6). The difference between the pegs and the floats, however, is not statistically significant at conventional levels (column 3), despite the large quantitative difference between the estimates. This can be explained by less-pronounced differences—and more heterogeneity—between the exchange-rate regimes in advanced countries, as the confidence intervals around the output responses in advanced pegs are particularly wide. Indeed, countries in a monetary union (e.g., the euro area) are pegs with respect to other countries in the union but floats relative to currencies

[14]While we verify that there is little qualitative difference between standard errors clustered by country and the Driscoll and Kraay (1998) standard errors in most cases, we note that the difference between the effects of positive and negative spending in developing countries (column 6) becomes insignificant with the Driscoll–Kraay errors, with a corresponding p-value of 0.113.

[15]Overall, these noisier estimates are in accord with the pattern that the multipliers are larger in recessions and for negative spending. For an example of a systematic approach to estimate multidimensional heterogeneity with more balanced data, see Suárez Serrato and Zidar (2018).

[16]The state-dependent multipliers for *positive* shocks are noisy, and thus we cannot reject either small or large values in each state. With the proportion of positive shocks at 40 percent and the incidence of recessions at 20 percent, less than 10 percent of observations correspond to positive shocks in recessions.

Table 6: Exchange-Rate Regimes and Trade Openness

	Full sample			Full sample			Open economies		
	Peg	Float	p-value	Closed	Open	p-value	Peg	Float	p-value
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
One year	1.196*	0.448		1.978***	0.366		1.443*	0.172	
	(0.681)	(0.579)	0.409	(0.752)	(0.541)	0.084	(0.856)	(0.435)	0.183
	[8.1]	[28.8]		[8.2]	[14.6]		[6.1]	[31.8]	
Two years	1.453**	0.653		1.702**	0.699		1.730**	0.481	
	(0.613)	(0.526)	0.349	(0.740)	(0.536)	0.275	(0.728)	(0.446)	0.149
	[13.9]	[27.0]		[7.6]	[18.3]		[9.5]	[25.5]	
Three years	1.172*	0.469		1.220*	0.514		1.286	0.428	
	(0.661)	(0.479)	0.372	(0.700)	(0.511)	0.417	(0.812)	(0.466)	0.310
	[6.6]	[30.8]		[4.9]	[16.4]		[4.4]	[30.7]	
Observations	1,250	1,649		1,278	1,711		856	813	

Notes: The exchange-rate classification in columns 1–3 and 7–9 is based on Klein and Shambaugh (2008), extended to the end of the sample. In columns 4–9, countries are classified as open to trade if their total trade (exports plus imports) is, on average, above 60 percent of GDP. p-values in columns 3, 6, and 9 are for the differences between the point estimates in the two preceding columns. Standard errors clustered by country are in parentheses. The first-stage effective F-statistics are in brackets. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

outside the union. The differences between the pegs and the floats in developing countries, most of which are not in a monetary union, are material.[17]

Our empirical estimates are consistent with theoretical models emphasizing the role of central-bank constraints under a fixed exchange rate (e.g., Mundell, 1963). In the models with free capital movement, a fiscal expansion is not offset by monetary tightening due to the central bank's commitment to maintain the peg, leading to a relatively strong response of output.[18] Our estimates are also consistent with Ilzetzki et al. (2013), who document a long-run multiplier of 1.4 for pegs and a multiplier statistically indistinguishable from zero for floats. Corsetti et al. (2012) and Born et al. (2013) also find larger output responses under a fixed exchange rate, but in samples that include only developed countries. We extend those studies' results on exchange-rate regimes to a more comprehensive sample of countries and a different identification strategy.

5.3.2. Trade Openness

We estimate the multipliers separately for open and closed economies. Following the literature, we define an economy as open if the share of total international trade (exports plus imports) is, on average, greater than 60 percent of GDP. The results in columns 4 through 6 of Table 6 show that fiscal policy is more effective in closed economies than in open economies. A $1 increase

[17]Similarly, one could conjecture that the fiscal multiplier may be relatively large when monetary policy is constrained due to other reasons, such as the zero lower bound.

[18]While, due to limited data, we do not include capital controls in the baseline specification, including capital controls in a sample with available data, as in Fernández et al. (2016), does not materially affect the estimates.

in government spending, on impact, leads to an increase in GDP of $1.98 in closed economies and $0.37 in open economies. The closed-economy multiplier is larger than the open-economy multiplier at all horizons considered, but the difference is statistically significant at the 10 percent level only at a one-year horizon.

The relative sizes of the multipliers in closed and open economies are, again, consistent with those in Ilzetzki et al. (2013). While, quantitatively, our estimates for closed economies are larger than the ones in that study, due to wide confidence intervals we cannot reject that the magnitudes are equal in the two studies. Our results are also consistent with the transmission mechanism documented by Miyamoto et al. (2019), who find that, in open economies, a fiscal expansion leads to a statistically significant appreciation of the domestic currency and a significant decrease in the current account, whereas in closed economies the appreciation is insignificant and the current-account decline is smaller. Hence, the stimulative effects of fiscal policy in open economies may spill over to foreign producers through a higher level of imports, while in closed economies the gains are concentrated in domestic production. However, we do not lend empirical support to mechanisms emphasizing positive effects of trade linkages on consumption, thereby giving rise to multipliers that are larger in open economies than in closed economies (e.g., Cacciatore and Traum, 2020).

5.3.3. Exchange Rates versus Trade Openness

Because the choice of an exchange-rate regime is more consequential for open economies, we estimate the multipliers for pegs and floats in the sample of open economies (columns 7–9 of Table 6). This exercise also helps us better understand the results just discussed, as it addresses the concern that open economies may choose systematically to have a different exchange-rate regime than in closed economies. We find that the multiplier is larger for pegs than for floats even when we focus on open economies only. The multiplier for open-economy pegs at 1.44 is slightly larger than in the full sample of pegs, and the multiplier for open-economy floats at 0.17 is somewhat smaller than for all floats. However, these differences from the full sample are small and not statistically significant. We also note that, in our sample, open economies do not favor one exchange-rate

regime over the other, as the episodes of pegs and floats are balanced.[19] Hence, our baseline results on the role of exchange-rate regimes hold for open economies and are not driven by trade openness.

Summary

Our empirical estimates indicate that the size of the fiscal multiplier can vary significantly. While our baseline estimates are consistent with evidence from different samples and methods reported in the literature, such similarities may be misleading, because the estimated effects of fiscal policy differ materially across countries and episodes in the sample. Focusing on potential sources of such heterogeneity, we document empirical support for the multipliers being relatively large in recessions, under exchange-rate pegs, and in closed economies. These findings lend empirical support to prominent macroeconomic models emphasizing each of these factors. We also provide new evidence on the relevance of military spending as an instrument for government spending in international data. Military spending appears to be a relevant instrument for government spending in the developing-countries sample, and it is particularly relevant in recessions.

6. Robustness

We now address concerns about our identification strategy and the robustness of our results. Our approach involves a battery of checks wherein countries and episodes with a suspected violation of the identifying assumption are dropped from the baseline sample.[20]

6.1. Financial Crises

We start by considering the effects of financial crises. Besides causing negative output effects, financial crises may prompt a decrease in military spending through budget cuts, leading to an omitted-variable bias. To check the extent of this issue, we drop observations with financial crises,

[19] However, only one-third of observations for closed economies correspond to pegs. This is not surprising because countries that do not rely significantly on the foreign sector may have limited incentives to intervene in the foreign exchange market. Moreover, some episodes of pegging identified by de facto classifications could be explained by limited volatility in the currency value due to stable fundamentals, and not by proactive policies to maintain a fixed exchange rate.

[20] Due to space constraints, we relegate the results of numerous other robustness checks to Online Appendix C.5. For example, the results on wars can be found in Table C.11. Estimates obtained for a sample of large shocks are in Table C.13. Results of the exercise where we remove observations with a negative correlation between military spending and total government spending are in Table C.14. And OLS estimates are in Table C.15.

Table 7: Robustness to Sample Composition

	No Financial Crises			No Autocracies			No Arms Importers		
	All	Adv.	Dev.	All	Adv.	Dev.	All	Adv.	Dev.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
One year	1.148***	1.843***	1.120**	1.009**	1.885**	0.899**	1.001**	1.697**	0.981*
	(0.432)	(0.692)	(0.469)	(0.394)	(0.721)	(0.443)	(0.479)	(0.722)	(0.496)
	[15.8]	[107.2]	[11.7]	[52.9]	[106.2]	[61.3]	[11.1]	[91.2]	[10.0]
Two years	1.101**	1.911**	1.060**	0.950**	2.196**	0.771*	1.049**	1.910**	0.990**
	(0.445)	(0.838)	(0.465)	(0.389)	(0.965)	(0.399)	(0.492)	(0.934)	(0.497)
	[15.4]	[61.8]	[13.2]	[38.4]	[83.8]	[80.8]	[19.8]	[88.9]	[18.7]
Three years	0.570	1.588**	0.514	0.705**	1.779**	0.536	0.722	1.550*	0.659
	(0.422)	(0.708)	(0.453)	(0.342)	(0.892)	(0.373)	(0.491)	(0.856)	(0.511)
	[7.6]	[56.8]	[6.3]	[28.0]	[72.0]	[40.3]	[15.0]	[73.3]	[13.2]
Observations	1,955	546	1,409	1,805	706	1,099	2,226	752	1,474

Notes: Standard errors clustered by country are in parentheses. The first-stage effective F-statistics are in brackets. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

as defined in Reinhart and Rogoff (2011). The multiplier estimates from this subsample are presented in columns 1 through 3 of Table 7. While the estimates for developing countries are slightly larger than the baseline estimates, they are quite similar to the baseline for advanced economies. In both subsamples, the new estimates are well within the baseline confidence intervals. We conclude that financial crises have only a minor effect on our results.

6.2. Anticipation of Military Shocks and Governance

Changes in military spending may be anticipated; some military programs are announced in advance and implemented over a longer period of time. The timing of the shock is therefore crucial for proper accounting of its effects (e.g., Ramey, 2011; Auerbach and Gorodnichenko, 2016). While we do not have data on military-spending announcements, using annual data allows us to overcome the mismatch between the news and implementation, as long as both take place within a calendar year. In addition, military spending on the programs initiated in the past and implemented over a long period are, in part, partialled out by the lags of spending used as controls.

To address the issue arising from longer-term programs, we examine the effects of governance and political instability on our results. A higher degree of political uncertainty or worsening in the governance practices may lead to an increase in expectations of military spending, due to potential civil unrest. Less-democratic governments may also be associated with a higher degree of political uncertainty and a higher probability of war, since it is easier for them to engage in hostilities, due to insufficient checks and balances. To this end, we employ the democracy and autocracy indices from the Polity IV Project. Each index varies from zero to ten, and we define a country as an autocracy in a given year if the democracy index is smaller than the autocracy index. We then

drop autocracies from the sample, and re-estimate the multipliers. Columns 4 through 6 of Table 7 report estimates qualitatively similar to the baseline.[21]

6.3. Arms Imports

Our next concern is arms imports. Most countries do not have significant arms production and therefore import arms. In countries where arms imports make up an especially large share of military spending, such spending may have a smaller effect on output due to a smaller increase in absorption. To understand how large arms importers differ from other countries in the sample, we exclude countries for which the average share of arms imports in total military spending is larger than 30 percent, which results in dropping more than 10 percent of the observations. As columns 7 through 9 of Table 7 suggest, excluding large arms importers does not materially affect our results.

6.4. Commodity Exports and Prices

The military-spending budget may be linked to commodity exports. In some cases, such a relation is legally binding. For example, Chile's Copper Law allocates 10 percent of copper export revenues to the military. In large commodity-exporting countries, an increase in commodity prices may therefore lead to an increase in GDP and in military spending. To address this concern, columns 1 through 3 of Table 8 exclude from our baseline sample large commodity exporters, which are defined as countries with average commodity exports above 50 percent of total exports. The results appear similar to our baseline. This conclusion also holds when we exclude large oil exporters.

In a separate exercise, we control for commodity prices. While large commodity exporters may increase their military spending when commodity prices are high and decrease it when they are low, commodity importers may exhibit the opposite pattern. Because oil is a major commodity that has been given disproportionate attention in the literature on supply shocks, we focus on oil prices. Oil is also a major commodity import in many countries with scarce natural resources. To control for oil prices, we use data on West Texas Intermediate crude oil. The estimates from this

[21] In another exercise, we use the World Bank's *Worldwide Governance Indicators* data, which contain a country's percentile rank of political stability and absence of violence and terrorism. We consider a country to be politically stable if it is ranked above the median. While the data are available for only a small subsample, controlling for political instability does not materially affect the multipliers in that sample.

Table 8: Commodity Exports and Prices

	No Commodity Exporters			Oil Prices			No Giant Oil Discoveries		
	All	Adv.	Dev.	All	Adv.	Dev.	All	Adv.	Dev.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
One year	0.684*	1.554**	0.684*	0.834**	1.679**	0.814**	1.303***	1.733**	1.299**
	(0.358)	(0.684)	(0.373)	(0.382)	(0.727)	(0.389)	(0.484)	(0.732)	(0.542)
	[26.2]	[94.9]	[31.2]	[15.2]	[91.3]	[14.3]	[19.6]	[72.5]	[16.0]
Two years	0.608	1.745**	0.562	0.861**	1.877**	0.818**	1.144**	1.853*	1.112**
	(0.367)	(0.852)	(0.375)	(0.394)	(0.933)	(0.393)	(0.457)	(0.952)	(0.497)
	[47.9]	[80.8]	[71.6]	[27.7]	[90.3]	[27.4]	[28.2]	[74.0]	[33.0]
Three years	0.337	1.381*	0.263	0.546	1.504*	0.499	0.679*	1.548*	0.637
	(0.348)	(0.745)	(0.365)	(0.378)	(0.842)	(0.387)	(0.400)	(0.841)	(0.447)
	[46.4]	[67.9]	[54.4]	[21.4]	[74.5]	[19.7]	[20.2]	[59.5]	[24.0]
Observations	1,851	708	1,143	2,563	774	1,789	1,780	519	1,261

Notes: Standard errors clustered by country are in parentheses. The first-stage effective F-statistics are in brackets. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

specification (columns 4–6 of Table 8) are very similar to our baseline estimates, in the pooled sample and in the advanced- and developing-countries subsamples.

Finally, giant oil discoveries have been discussed in the literature as another factor that affects both military spending and output. For example, Lei and Michaels (2014) find that giant oil discoveries increase the probability of armed conflict. As a robustness check, we exclude large oil discoveries identified in that study. The resulting estimates are again comparable to the baseline, with slightly larger (than in the baseline) multipliers in developing countries (columns 7–9 of Table 8). This conclusion also holds when we drop from the sample countries with at least one giant oil discovery during the sample period.

7. Conclusion

Using data on military spending for more than a hundred countries, we estimate an annual government-spending multiplier, pooled across the sample, in the range 0.74–0.83 at a one-year horizon. The multiplier estimates remain significant over longer horizons. We also find that the multiplier is larger in developed countries than in developing countries, in recessions than in expansions, under a fixed exchange rate than under a floating regime, and in closed economies than in open economies. Hence, there is a wide range of economic conditions for which fiscal policy can be particularly powerful.

Our findings have important implications for policymakers, as the decisions on whether and how to use fiscal policy should account for particular economic circumstances and environments. Our findings also have implications for the identification of government-spending shocks. We find

that the military-spending approach produces multipliers comparable to those from other methods, even when extended to a broad sample of developing countries, and that they are qualitatively robust to tests involving various controls and samples used to address potential violations of the identifying restrictions. However, we acknowledge that, reflecting the underlying heterogeneity, the confidence bands of the multipliers are wide relative to the estimates obtained for a single country and for a small number of countries with similar economic institutions.

Our work also touches on some important issues that need further investigation but require additional data.[22] While we do not find a strong response of tax rates to government-spending shocks for a subset of countries with available data, we hope that tax-financed multipliers will be given further assessment in future work. Furthermore, we find only mild evidence of coordination between the fiscal and monetary authorities, which may point to some institutional constraints. It remains a question if such coordination can affect the multiplier size in a quantitatively important manner. Finally, the role of spending composition as well as other aspects of fiscal programs can be explored further when more granular data become available to researchers.

Acknowledgements

For comments and suggestions, we thank Olivier Coibion, Bill Dupor, Andy Glover, Yuriy Gorodnichenko, Josh Hausman, Wataru Miyamoto, Thuy Lan Nguyen, Maurice Obstfeld, Christina Romer, David Romer, Matthew Shapiro, Juan Carlos Suárez Serrato, Christina Wang, Johannes Wieland, Randy Wright, as well as the anonymous referees. We also thank participants at seminars and conferences where we presented this research. Nikhil Rao provided superb research assistance. An earlier version of this paper has circulated under the title "Output Response to Government Spending: Evidence from New International Military Spending Data." The views expressed herein are those of the authors and are not necessarily those of the Federal Reserve Bank of Boston or the Federal Reserve System. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References

Aiyagari, S. R., Christiano, L. J., Eichenbaum, M., 1992. The Output, Employment, and Interest Rate Effects of Government Consumption. Journal of Monetary Economics 30 (1), 73–86.

Alesina, A., Ardagna, S., 1998. Tales of Fiscal Adjustment. Economic Policy 13 (27), 487–545.

Auerbach, A. J., Gorodnichenko, Y., 2012. Measuring the Output Responses to Fiscal Policy. American Economic Journal: Economic Policy 4 (2), 1–27.

[22] We discuss these issues in more detail, using data from a smaller sample of countries, in Online Appendix D.

Auerbach, A. J., Gorodnichenko, Y., 2013. Fiscal Multipliers in Recession and Expansion. In: Alesina, A., Giavazzi, F. (Eds.), Fiscal Policy after the Financial Crisis. University of Chicago Press, Chicago, pp. 63–98.

Auerbach, A. J., Gorodnichenko, Y., 2016. Effects of Fiscal Shocks in a Globalized World. IMF Economic Review 64 (1), 177–215.

Barnichon, R., Debortoli, D., Matthes, C., 2021. Understanding the Size of the Government Spending Multiplier: It's in the Sign. Review of Economic Studies. Forthcoming.

Barro, R. J., 1981. Output Effects of Government Purchases. Journal of Political Economy 89 (6), 1086–1121.

Barro, R. J., Redlick, C. J., 2011. Macroeconomic Effects from Government Purchases and Taxes. Quarterly Journal of Economics 126 (1), 51–102.

Beetsma, R., Giuliodori, M., 2011. The Effects of Government Purchases Shocks: Review and Estimates for the EU. Economic Journal 121 (550), 4–32.

Blanchard, O., Perotti, R., 2002. An Empirical Characterization of the Dynamic Effects of Changes in Government Spending and Taxes on Output. Quarterly Journal of Economics 117 (4), 1329–1368.

Born, B., Juessen, F., Müller, G. J., 2013. Exchange Rate Regimes and Fiscal Multipliers. Journal of Economic Dynamics and Control 37 (2), 446–465.

Bräuning, F., Sheremirov, V., 2019. Output Spillovers from U.S. Monetary Policy: The Role of International Trade and Financial Linkages. Working paper no. 19-15, Federal Reserve Bank of Boston.

Cacciatore, M., Traum, N., 2020. Trade Flows and Fiscal Multipliers. Review of Economics and Statistics. Forthcoming.

Chodorow-Reich, G., 2018. Geographic Cross-Sectional Fiscal Spending Multipliers: What Have We Learned? American Economic Journal: Economic Policy 11 (2), 1–34.

Christiano, L., Eichenbaum, M., Rebelo, S., 2011. When Is the Government Spending Multiplier Large? Journal of Political Economy 119 (1), 78–121.

Collier, P., 2006. War and Military Expenditure in Developing Countries and Their Consequences for Development. Economics of Peace and Security Journal 1 (1), 10–13.

Corsetti, G., Meier, A., Müller, G. J., 2012. What Determines Government Spending Multipliers? Economic Policy 27 (72), 521–565.

Driscoll, J. C., Kraay, A. C., 1998. Consistent Covariance Matrix Estimation with Spatially Dependent Panel Data. Review of Economics and Statistics 80 (4), 549–560.

Dupor, B., Guerrero, R., 2017. Local and Aggregate Fiscal Policy Multipliers. Journal of Monetary Economics 92, 16–30.

Fernández, A., Klein, M. W., Rebucci, A., Schindler, M., Uribe, M., 2016. Capital Control Measures: A New Dataset. IMF Economic Review 64 (3), 548–574.

Galí, J., López-Salido, J. D., Vallés, J., 2007. Understanding the Effects of Government Spending on Consumption. Journal of the European Economic Association 5 (1), 227–270.

Gavin, M., Perotti, R., 1997. Fiscal Policy in Latin America. In: Bernanke, B. S., Rotemberg, J. J. (Eds.), NBER Macroeconomics Annual 1997. Vol. 12. MIT Press, Cambridge, MA, pp. 11–72.

Gleditsch, N., Wallensteen, P., Eriksson, M., Sollenberg, M., Strand, H., 2002. Armed conflict 1946–2001: A new dataset. Journal of Peace Research 39 (5), 615–637.

Gordon, R. J., Krenn, R., 2010. The End of the Great Depression 1939–41: Policy Contributions and Fiscal Multipliers. Working paper no. 16380, National Bureau of Economic Research.

Hagedorn, M., Manovskii, I., Mitman, K., 2019. The Fiscal Multiplier. Working paper no. 25571, National Bureau of Economic Research.

Hall, R. E., 1986. The Role of Consumption in Economic Fluctuations. In: Gordon, R. J. (Ed.), The American Business Cycle: Continuity and Change. University of Chicago Press, Chicago, pp. 237–266.

Hall, R. E., 2009. By How Much Does GDP Rise If the Government Buys More Output? Brookings Papers on Economic Activity 40 (2), 183–249.

Ilzetzki, E. O., Mendoza, E. G., Végh, C. A., 2013. How Big (Small?) Are Fiscal Multipliers? Journal of Monetary Economics 60 (2), 239–254.

Jordà, O., 2005. Estimation and Inference of Impulse Responses by Local Projections. American Economic Review 95 (1), 161–182.

Klein, M. W., Shambaugh, J. C., 2008. The Dynamics of Exchange Rate Regimes: Fixes, Floats, and Flips. Journal of International Economics 75 (1), 70–92.

Kraay, A., 2012. How Large Is the Government Spending Multiplier? Evidence from World Bank Lending. Quarterly Journal of Economics 127 (2), 829–887.

Kraay, A., 2014. Government Spending Multipliers in Developing Countries: Evidence from Lending by Official Creditors. American Economic Journal: Macroeconomics 6 (4), 170–208.

Lei, Y.-H., Michaels, G., 2014. Do Giant Oilfield Discoveries Fuel Internal Armed Conflicts? Journal of Development Economics 110, 139–157.

Leigh, D., Devries, P., Freedman, C., Guajardo, J., Laxton, D., Pescatori, A., 2010. Will It Hurt? Macroeconomic Effects of Fiscal Consolidation. In: World Economic Outlook: Recovery, Risk, and Rebalancing. International Monetary Fund, Washington, pp. 93–124.

Michaillat, P., 2014. A Theory of Countercyclical Government Multiplier. American Economic Journal: Macroeconomics 6 (1), 190–217.

Miyamoto, W., Nguyen, T. L., Sheremirov, V., 2019. The Effects of Government Spending on Real Exchange Rates: Evidence from Military Spending Panel Data. Journal of International Economics 116, 144–157.

Montiel Olea, J. L., Pflueger, C., 2013. A Robust Test for Weak Instruments. Journal of Business & Economic Statistics 31 (3), 358–369.

Mundell, R. A., 1963. Capital Mobility and Stabilization Policy under Fixed and Flexible Exchange Rates. Canadian Journal of Economics and Political Science 29 (4), 475–485.

Nakamura, E., Steinsson, J., 2014. Fiscal Stimulus in a Monetary Union: Evidence from US Regions. American Economic Review 104 (3), 753–792.

Parker, J. A., 2011. On Measuring the Effects of Fiscal Policy in Recessions. Journal of Economic Literature 49 (3), 703–718.

Perotti, R., 2008. In Search of the Transmission Mechanism of Fiscal Policy. In: Acemoglu, D., Rogoff, K. S., Woodford, M. (Eds.), NBER Macroeconomics Annual 2007. Vol. 22. University of Chicago Press, Chicago, pp. 169–226.

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

Ramey, V. A., Zubairy, S., 2018. Government Spending Multipliers in Good Times and in Bad: Evidence from US Historical Data. Journal of Political Economy 126 (2), 850–901.

Reinhart, C., Rogoff, K., 2011. From Financial Crash to Debt Crisis. American Economic Review 101 (5), 1676–1706.

Rotemberg, J. J., Woodford, M., 1992. Oligopolistic Pricing and the Effects of Aggregate Demand on Economic Activity. Journal of Political Economy 100 (6), 1153–1207.

Sarkees, M. R., Wayman, F., 2010. Resort to War: 1816–2007. CQ Press, Washington.

Shoag, D., 2016. The Impact of Government Spending Shocks: Evidence on the Multiplier from State Pension Plan Returns. Working paper, Harvard University.

Staiger, D., Stock, J. H., 1997. Instrumental Variables Regression with Weak Instruments. Econometrica 65 (3), 557–586.

Stålenheim, P., Perdomo, C., Sköns, E., 2008. Military Expenditure. In: SIPRI Yearbook 2008: Armaments, Disarmament and International Security. Oxford University Press, New York, Ch. 10, pp. 175–206.

Suárez Serrato, J. C., Wingender, P., 2014. Estimating the Incidence of Government Spending. Working paper, Duke University.

Suárez Serrato, J. C., Wingender, P., 2016. Estimating Local Fiscal Multipliers. Working paper no. 22425, National Bureau of Economic Research.

Suárez Serrato, J. C., Zidar, O., 2018. The Structure of State Corporate Taxation and Its Impact on State Tax Revenues and Economic Activity. Journal of Public Economics 167 (C), 158–176.

Talvi, E., Végh, C. A., 2005. Tax Base Variability and Procyclical Fiscal Policy in Developing Countries. Journal of Development Economics 78 (1), 156–190.

Woodford, M., 2011. Simple Analytics of the Government Expenditure Multiplier. American Economic Journal: Macroeconomics 3 (1), 1–35.