Fiscal Multipliers of Different Government Spending Categories

Thomas van Gemert*, Lenard Lieb and Tania Treibich

Maastricht University

February 2019

Abstract

In this paper we estimate multipliers of different categories of government spending in the United States. We estimate short-horizon and long-horizon multipliers for defense spending and categories of non-defense spending. We offer a new way to capture non-defense spending with intergovernmental transfers and propose a projection-based approach to study long horizons. By using a panel moving blocks bootstrap, we test whether multipliers of different categories of government spending differ significantly. We find that defense spending results in a lower multiplier than non-defense spending.

Keywords: Fiscal Policy, Government Spending, Local Fiscal Multiplier

JEL Codes: E62, H30, H50, H70, H71, H72, H77, N42

A central question in macroeconomics is whether the government can effectively stimulate the economy by increasing spending or decreasing taxes. Previous studies have estimated the aggregate fiscal multiplier at the national level using time series data (Blanchard and Perotti, 2002; Barro and Redlick, 2011; Ramey, 2011; Ramey and Zubairy, 2018) or at the regional level using cross-sectional data (Nakamura and Steinsson, 2014; Dupor and Guerrero, 2017).[1] A clear advantage of the cross-sectional (or panel) approach is that this allows

*Maastricht University, School of Business and Economics. E-mail: t.vangemert@maastrichtuniversity.nl

[1]It is also possible to take a multi-country approach for the cross-sectional analysis (Ilzetzki et al., 2013).

to control for the aggregate business cycle through year-fixed effects. Hence, fewer variables are necessary to control for confounding business cycle variables when estimating the relation between government spending and income. Furthermore, by adding the cross-sectional dimension the number of observations increases sharply, which makes estimation more feasible when the time dimension is short. Authors that followed the cross-sectional approach used specific categories of government spending, such as defense spending (Nakamura and Steinsson, 2014; Dupor and Guerrero, 2017) or non-defense spending (Clemens and Miran, 2012; Shoag, 2016; Chodorow-Reich, 2017), because of methodological considerations. However, differences in methodology cannot fully explain the difference between estimates. There are economic reasons why categories of spending result in distinct effects. No study has provided a structured investigation of how the the fiscal multiplier is dependent on government spending categories. The aim of this paper is to provide comparable estimates for different categories of U.S. federal government spending.

This paper contributes to the empirical literature on the local fiscal multiplier. We follow the cross-sectional approach to study the effect of federal defense and non-defense spending, by taking into account the specific structure of government spending. Defense spending is the only category of federal government spending in the U.S. without any involvement of state governments. We use this category as a direct federal spending stimulus to state economies. For non-defense spending categories, we propose to use intergovernmental transfers from the federal to the state government. The categories that we consider are: education, health care, highways, and Medicaid (public welfare). For all spending categories we estimate both a static short-horizon multiplier and dynamic long-horizon multipliers. In the static analysis, we use instrumental variable (IV) estimation, in line with many studies the related literature. For the dynamic analysis, we use a projection-based approach, because the instruments are not suitable to use for longer horizons. The projection-based approach makes it possible to estimate dynamic impulse responses of both government spending and personal income, following the local projections set-up by Jordà (2005). Furthermore, we implement a novel way to construct inference on the point estimates for the fiscal multiplier by using the moving blocks bootstrap (Gonçalves, 2011). These confidence intervals allow to test whether there is a significant difference between multiplier estimates.

In section 1 we provide an overview of the related literature. In section 2 we discuss the data that we use in the static and dynamic analysis. We present our methodology and the corresponding results for the static and dynamic analysis in section 3 and 4 respectively. Finally, we conclude paper in section 5 and give some directions for future research.

1 Related literature

The estimated multiplier in the cross-sectional literature is often labeled as the 'local fiscal multiplier', because sub-national data are used. Nakamura and Steinsson (2014) show that the estimated local fiscal multiplier should be interpreted as an open economy relative multiplier, which is different from the closed economy aggregate multiplier that is estimated with time series data (Blanchard and Perotti, 2002; Barro and Redlick, 2011; Ramey, 2011; Ramey and Zubairy, 2018). Estimating the local fiscal multiplier is challenging, because government spending at the state level is endogenous due to two sources. First, there is two-way causality between government spending and business cycle fluctuations. Since tax revenues are positively correlated to income, government spending is also correlated to income. If an increase in government spending stimulates income, then the direction of causality cannot be observed. This problem arises both with time series and panel data. Second, there is a specific problem when using state-level data. Many spending programs by the federal government intend to provide economic support to states. Hence, the federal government spending stimulus to a state is endogenous to the state-level business cycle. The multiplier estimates are biased upwards, because multipliers are higher in economic downturns (Auerbach and Gorodnichenko, 2012).

To solve endogeneity problems, the common approach in the literature is to use instrumental variables. Many instruments used in the literature apply to a specific spending category at a particular governmental level, i.e. federal, state or local. In a fiscal union, multiple levels of government are often engaged. An exception is defense spending, which only takes place at the federal level. Other categories of spending are conducted by the federal government in cooperation with state and local governments, which creates fiscal flows between governments. These flows are summarized in the non-hierarchical overview in figure 1 below. We summarize the existing literature in this section based on this classification.

Figure 1: Fiscal policy flows in a fiscal union

$$
\begin{tikzcd}
\text{Federal Government} \arrow[d, "\text{B1}"] \arrow[dd, bend right=60, "\text{A}"'] \\
\text{State Government} \arrow[d, "\text{B2}"] \arrow[d, shift left=8, "\text{C}"] \\
\text{State Economy}
\end{tikzcd}
$$

First, there are papers that investigated the effect of direct federal government spending (channel A) on income in a state or region. Existing studies used defense spending as exogenous variable to instrument federal spending at the state or local level. Both Nakamura and Steinsson (2014) and Dupor and Guerrero (2017) used instrumental variables (IV) to exploit the heterogeneity in the response of state-level defense contracts to national variation in defense spending. The identifying assumption is that the federal government does not increase national defense spending to give a disproportional economic stimulus, because this state is doing economically poorly compared to other states. Both studies used this approach by Bartik (1991) to instrument military spending per state with the change in national defense spending, scaled by the average level of military spending in that state to its output in previous years. Nakamura and Steinsson (2014) estimated a local fiscal multiplier of 1.5 for a two-year horizon, but Dupor and Guerrero (2017) showed that the multiplier is probably lower, depending on the estimation period. Next to these studies, two papers used historical policy examples from the 1930s to illustrate the effect of direct federal spending at the local level. Biolsi (2015) focused on naval shipbuilding and Hausman (2016) used veterans bonuses. These studies did not provide estimates for the local fiscal multiplier.

Other studies focused on the effect of indirect federal (non-defense) spending at the state level (channel B1 and B2).[2] The federal government provides state governments with grants to carry out a particular spending program. The effectiveness of an indirect federal fiscal shock depends on how the shock is transmitted by state governments and whether the grants (channel B1) crowd out spending by state governments (channel B2 and C). To avoid this, the federal government attaches conditions to grants. In total, 77% of granted amount involves either matching or maintenance-of-effort requirements (FFIS, 2016). Matching requirements require a state government to use own revenues to match the amount granted by the federal government. Maintenance-of-effort requirements require state governments to not cut spending in the program tied to the grant when it receives federal funds. Empirical studies have found that federal grants do not crowd out state spending, which is the so-called flypaper-effect (Gamkar and Oates, 1996; Nesbit and Kreft, 2009).[3]

Many papers that used indirect federal spending tried to find an exogenous driver of the grants to state governments, such as specific rules in grant mechanisms (Chodorow-Reich et al., 2012; Wilson, 2012; Fishback and Kachanovskaya, 2015; Suárez Serrato and Wingender, 2016; Chodorow-Reich, 2017). These rules are used to instrument grants from the federal to the state government (channel B1). Alternatively, Leduc and Wilson (2013) did not use an IV-set up, but used the difference between the expected amount of grants and the actual amount received to estimate the indirect federal government spending shock. All fiscal multiplier estimates for this spending channel range between 1.4 and 2.0.[4]

Finally, there are papers that used internally financed state government (non-defense) spending (channel C).[5] In this case state governments finance the spending with state taxation, deficits, or other funds, such as rainy-day funds, pension funds, or lottery funds. Clemens and Miran (2012) analysed deficit-financed spending by state governments. They used differences in state balanced-budget requirements and estimated a multiplier of 0.3. Shoag (2016) analyzed an external shock to state government spending by exploiting windfalls in state pension funds and estimated a multiplier close to 2.0.

[2] Suárez Serrato and Wingender (2016) did not clearly distinguish between direct and indirect spending.
[3] Specific categories are analyzed by for example Gordon (2004) (education) or Knight (2002) (highways).
[4] Except Fishback and Kachanovskaya (2015), but this study used data from the Great Depression.
[5] Adelino et al. (2017) analyzed the effect of municipal spending.

2 Data

The objective of this paper is to estimate the effect of federal government spending categories on income. We use panel data for 50 U.S. states to exploit the cross-sectional differences between states. We use data on personal income at the state level. These data are available in the Regional Economic Accounts from the U.S. Bureau of Economic Analysis. The observations range from 1963 to 2014 for all 50 U.S. states. Personal income measures income of all residents, including income received from other states. We interpret the level of personal income as a measure for the economic well-begin of residents in one state relative to other states. The growth rate of income captures the change in the welfare of residents in a particular state. The variable includes wages, proprietors' income, dividends, interest, rents and government benefits. We use these data in real per capita, using the population estimate and deflator from Dupor and Guerrero (2017). In line with Shoag (2016), we use personal income instead of output, due to limitations in data availability.[6] However, personal income follows fluctuations of output closely.

Next, we also use data on different categories of government spending.[7] One of these categories is defense spending, which is the only direct federal spending category, because there is no involvement from the state government (channel A). To measure defense spending at the state level, we use data from Dupor and Guerrero (2017) on state-level defense contracts. This database is the updated version of the initial work from Nakamura and Steinsson (2014). The data are available from 1963 until 2014 for all U.S. states. The database has been constructed using annual reports on defense contracts. The included contracts are manufacturing (processing and assembling), construction, and service contracts (transportation and communication). The amounts are aggregated to obtain an estimate for military procurement actions per state. We use the variable in real per capita. It captures the total real amount spent on defense procurement by the federal government in a particular state.

There are two possible concerns when using this variable. First, it is possible that con-

[6] There is a break in 1997, because the BEA shifted industry classification from SIC to NAICS.

[7] Strictly speaking, we use proxies of government spending categories, because the variables that we use do not measure the categories of government spending entirely.

tractors rely on sub-contractors in other states. However, Nakamura and Steinsson (2014) have shown that controlling for subcontracting does not change the results. Second, it is a problem when the actual work by firms does not take place in the same year as the announcement (i.e. the signing) of the contract. Since households and firms are forward looking, they start anticipating on the future effect of a federal spending stimulus. This time inconsistency when estimating the spending shock could create biased estimates. Dupor and Guerrero (2017) have shown that their results are robust to anticipation by using the Ramey (2011) defense news series.

Many categories of non-defense spending are not carried out only by the federal government. Instead, the federal government cooperates with lower level governments by financing specific programs through intergovernmental transfers[8]. We use intergovernmental transfers from the federal government to state governments to capture channel B1. Examples of such grants are Medicaid or the Federal-Aid Highway Program. We use data on federal transfers by the U.S. Census bureau from the Annual Survey of State and Local Government Finances. We use both the aggregate measure and the individual categories: education, health, highways, and public welfare (Medicaid). In the related literature, studies often focus on specific grants, such as ARRA grants (Chodorow-Reich, 2017) or highway grants (Leduc and Wilson, 2013). The variable intergovernmental transfers consists off all categorical and block grants awarded both on a competitive basis or through formulas.[9] These grants are an important source of revenue for state governments. The FFIS (2018) notes that state governments finance 30% of their spending with intergovernmental transfers.

Data on intergovernmental transfers are available from 1963 to 2014 for all states. We transform the variable in real per capita. We focus on intergovernmental transfers, because these capture (indirect) federal non-defense spending in the most complete sense. A possible concern is that intergovernmental transfers measures revenue for state governments, but that these are not one-to-one linked with state government spending. Funds from the federal government could crowd out other parts of the state government budget, such as own-

[8]We will use the terms intergovernmental grants or transfers interchangably.
[9]94% of the amount of grants is formula based FFIS (2018)

revenue financed spending (channel C). However, 77% of the total grant amount from the federal government to state government contains either matching or maintenance-of-effort requirements (FFIS, 2016). It is therefore unlikely that grants crowd out state government spending. Hence, there is a close link between intergovernmental grants and state spending.

For non-defense spending, we use spending by the federal government through state governments.[10] We therefore use a measure for spending by the state government, which is financed by federal funds (channel B2). We use data by the U.S. Census Bureau from the Annual Survey of State and Local Government Finances on total state government spending and direct spending on the categories education, health, highways, and public welfare (Medicaid). These data are available from 1963 to 2014 for all states. We transform the variable in real per capita.

We include the following set of control variables in our dynamic analysis. We use the state unemployment rate to capture the state-specific business cycle in the dynamic analysis. Both personal income and government spending per state are positively correlated to the state business cycle. Ignoring this would create upward bias in the multiplier estimates. The source of the unemployment rate is the Bureau of Labor Statistics (BLS), Local Area Unemployment Statistics (LAUS). These data are available from 1976 until 2014.

To control for the dynamics in state government finances, we use the deficit ratio for the state government as control variable. State governments can run deficits, but many states are constrained by balanced budget requirements. These requirements affect the responsiveness of the state government budget. State government spending becomes more pro-cyclical, because it follows closer the revenues, which are correlated to business cycle fluctuations. Hence, if we would not include the deficit ratio, the multiplier estimates could be biased upward. We calculate the deficit ratio from state expenditures and revenues data by the U.S. Census Bureau.

Finally, we use political control variables on the state-level in the static analysis to control for dominance of political parties in states. When a state is dominated by one political party,

[10]Technically, this means that when we compare defense and non-defense spending, we use two different spending channels. This is unavoidable due to the specific structure of U.S government spending.

it can be easier for the party to gain influence on the national level, especially if that party also controls the federal administration. Ignoring this would also create upward bias. We therefore use data on the political dominance of the Democratic Party versus the Republican Party in the state senate, state house, and governorship. We updated the data by Klarner (2013) for the last 4 years using the State Partisan Composition tables from NCSL. [11]

3 Static analysis

The aim of the static analysis is to estimate the effect of different categories of federal government spending for short horizons. We compare the effect of defense versus non-defense spending and several categories of non-defense spending. We use instrumental variables (IV) estimation to measure the short run effect of a change to the different categories of government spending. We cannot use ordinary least squares (OLS), because of two endogeneity problems: two-way causality and ommitted variable bias. The IV approach is used by many related studies (Nakamura and Steinsson, 2014; Dupor and Guerrero, 2017; Shoag, 2016; Chodorow-Reich, 2017).

The validity of IV depends on two conditions. The first condition is about the relevance of the instrument. The instrument should be a strong instrument, i.e. the instrument should have significant explanatory power in the first-stage regression for the government spending category. This condition is testable with a weak instrument partial F-test. The second condition is the exclusion restriction. This condition states that the instrument should not directly affect the outcome variable, i.e. personal income. The effect from the instrument on the outcome variable should only go through government spending. Unfortunately, this condition is not testable. However, we argue that the use of the instruments is justified.

[11] Source: `www.ncsl.org/research/about-state-legislatures/partisan-composition.aspx`

3.1 Methodology

The method proceeds in three steps. First, we create the instruments for different government spending categories. Second, we run the first-stage regression for the government spending categories. Third, we estimate the second stage regression for personal income.

We start by creating the instruments. We follow Nakamura and Steinsson (2014) and Dupor and Guerrero (2017) in using the Bartik (1991) approach to construct the instruments. The intuition behind this method is that fluctuations of the national level of government spending are exogenous to state-specific business cycles. The federal government decides on the national level of spending without considering how this works out for specific states. This approach can be used to construct a state-specific exogenous variable. For this variable to be exogenous to the contemporaneous state business cycle, we should only use information from the federal level and state-specific characteristics from previous years.

We construct these instruments $Z_{i,t}$ by interacting the national one-period growth of a category of federal spending, multiplied with a state-specific scaling factor. Formally, this can be written down as follows, where we use for $X^n_{i,t}$ the national level of federal spending:

$$Z_{i,t} = \frac{s^X_{i,t}}{s^Y_{i,t}} \frac{X^n_t - X^n_{t-1}}{X^n_{t-1}} \tag{1}$$

This scaling factor is the lagged share of the state in the total national amount for that type of spending $s^X_{i,t}$, divided by the share in national personal income $s^Y_{i,t}$. Both shares are computed using the average of the previous year and the year before. This approach exploits the heterogeneity in state-level responses to national changes of government spending. The identifying assumption is that the federal government does not change federal spending, because states receiving a disproportionate amount of spending are doing relatively poorly.

This approach is used by Nakamura and Steinsson (2014) and Dupor and Guerrero (2017) for defense spending. However, the question is whether this also works for non-defense spending. Defense spending is mainly driven by international political events according to Ramey (2011), but this is not the case for non-defense spending. However, for non-defense spending we use transfers from the federal government to state governments (channel B1). From the perspective of the state government, the spending decision is external. The instrument is constructed such that it is exogenous to state-level business cycles. However, we should also control for the national business cycle, because non-defense spending is on a national level not only driven by exogenous events. We control for the national business cycle by including year-fixed effects in the regression.

The estimation strategy that we use follows a two-stage least squares (2SLS) set-up. We jointly estimate the first and second stage and compute directly the standard errors that are asymptotically robust to heteroskedasticity and autocorrelation. We use the same regression framework as Nakamura and Steinsson (2014). The first-stage regression looks as follows, where $G_{i,t}$ is either defense spending or state-government spending (for non-defense spending) and this is instrumented using the constructed instrument $Z_{i,t}$, and state and year fixed effects (μ_i and ν_t):

$$\frac{G_{i,t} - G_{i,t-2}}{Y_{i,t-2}} = \mu_i + \nu_t + \theta Z_{i,t} + \epsilon_{i,t} \tag{2}$$

The year fixed effects in this regression control for the contemporaneous correlation of the dependent and independent variable with the national business cycle.

According to the conditions for IV estimation, the instruments used in this first stage regression should be relevant (condition 1). We therefore report the F-statistic of the instrument exclusion test, i.e. the weak instrument test. A deficiency of this estimation set-up is that we instrument the two-year growth rate of government spending, but the instrument is constructed using the one-year growth rate. Although this approach is in line with the analysis of Nakamura and Steinsson (2014), we will later check the robustness of this approach by using the two-year growth rate when constructing the instrument.

We use the instrumented value in the second stage to estimate the effect on the two-year growth rate of personal income. The second stage regression can be written down as follows:

$$\frac{Y_{i,t} - Y_{i,t-2}}{Y_{t-2}} = \alpha_i + \gamma_t + \beta \frac{G_{i,t} - G_{i,t-2}}{Y_{i,t-2}} + \varepsilon_{i,t} \tag{3}$$

Here $Y_{i,t}$ is state-level personal income and $G_{i,t}$ is the instrumented value from the first stage of either defense contracts or state government spending for non-defense spending. This regression also includes state and year-fixed effects (α_i and γ_t). We follow the model specified by Nakamura and Steinsson (2014). This means that we include only time and year-fixed effects in the IV-regressions to control for the national business cycle.

This approach allows to compute directly the short-horizon multiplier, which is the coefficient β for either defense contracts or categories of state government spending in the second stage. However, since we use the two-year growth rate of income as dependent variable, we ignore the effect for longer horizons. We perform a dynamic analysis in section 4, where we analyze the effects for longer horizons.

According to the conditions for IV estimation, the instrument should only affect personal income indirectly through the government spending variables (condition 2). We can argue that this is the case. We use intergovernmental transfers for non-defense spending. These transfers are by construction intended to change government spending. Moreover, it is unlikely that the federal government changes the total national amount of defense spending or intergovernmental transfers to give a particular state directly a stimulus. On the other hand, defense spending is on the national level mainly driven by exogenous events.

3.2 Results

First, we provide the estimation results for defense spending versus non-defense spending. We investigate whether there is a difference in the short-horizon multiplier for these two categories of spending. For non-defense spending, we use the total amount of intergovernmental transfers to state governments to instrument total state government spending. The estimation results are shown in table 1 below.

The table shows in columns (1) and (3) the estimated multipliers from the second stage regressions. We estimate a 2-year multiplier around 1.2 for defense spending with an F-statistic[12] of 8.48. This result is close to the estimate from Nakamura and Steinsson (2014),

[12] This is an exlusion test for removing only the instrument in the first stage regression, while keeping the other control variables.

Table 1: Static results for defense versus non-defense spending

	(1) Pers. Income	(2) Pers. Income	(3) Pers. Income	(4) Pers. Income
Defense Spending	1.1659**	1.2291**	-	-
	(0.5069)	(0.5243)		
Non-defense Spending	-	-	2.0582	2.0166
			(1.4882)	(1.5407)
Fixed effects	State, Year	State, Year	State, Year	State, Year
Political controls	No	Yes	No	Yes
First stage F-stat.	8.48	8.23	4.04	3.64

Robust SEs in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

although we use a longer period for estimation. For non-defense spending we estimate a 2-year multiplier of 2.1 with an F-statistic of 4.04. However, the large standard errors prevent us from claiming that there is a significant difference between the estimated multipliers.

The first-stage coefficients indicate that both instruments suffer from the weak instrument problem. It is impossible to investigate this further, because we cannot add more control variables. It is hard to find relevant control variable that are exogenous to personal income. Including exogenous national variables is not an option, because these are absorbed in the year fixed effects. The results are also sensitive to the choice of the dependent variable. Using the one-year or three-year growth rate of personal income changes the estimated multiplier substantially. However, we show in section 4 that in a dynamic regression with more control variables a qualitatively similar result can be obtained.

We perform the following robustness checks. First, we add political control variables to control for the possibility that states receive a higher proportion of federal government spending, because a political party dominates the state legislation. Ignoring this would create upward bias in the multiplier estimates in the second stage. We therefore include dummies for the dominance of the Democratic Party in the state senate, state house and the state governorship. The results are shown in columns (2) and (4) in table 1. We find no significant change in the estimates when we include these control variables. We also check whether the results are sensitive to the horizon that we use when creating the instrument. It would make sense to create the instrument at the same horizon as the dependent variable in the first and second stage. When we use the two-year growth rate, the results do not change qualitatively (see appendix A.1), but the significance of the instrument for non-defense spending drops.

Second, we investigate the effect of different categories of non-defense spending. Our objective is to test whether there is a difference in the short-horizon multiplier for different categories of non-defense spending. The estimation results are shown in table 2 below. The first-stage F-statistics for education and health are very low, which means that these instruments are weak. The F-statistics highways and Medicaid are reasonable, but the estimated multipliers are unrealistic. The estimated multiplier for highways is negative and the multiplier for Medicaid is too high. Hence, we think that we should be cautious when using this method and interpreting the results. It is very likely that all instruments are bad instruments. We will investigate this further in section 4.

Table 2: Static results for categories of non-defense spending

	(1) Pers. Income	(2) Pers. Income	(3) Pers. Income	(4) Pers. Income
Education Spending	18.1166	-	-	-
	(92.1627)			
Highway Spending	-	-3.8842	-	-
		(2.6858)		
Health Spending	-	-	125.4364	-
			(288.6838)	
Medicaid Spending	-	-	-	9.2335**
				(3.7654)
Fixed effects	State, Year	State, Year	State, Year	State, Year
Political controls	No	No	No	No
First stage F-stat.	0.03	9.67	0.26	10.05
N	2450	2450	2450	2450

Robust SEs in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

To check the robustness of the results, we perform the same robustness checks as before. First, we include dummies to control for the dominance of the Democratic party in a state (see appendix A.2). This does not lead to a change in the results. Second, we use the two-year growth rate in the national government spending variables when creating the instruments (see appendix A.1). The estimated multipliers are still unrealistic when using this method.

4 Dynamic analysis

Next to the static analysis, which only shows the short-run impact of defense versus non-defense spending, we are also interested in the (cumulative) effect for longer horizons. The related literature has largely ignored these longer horizons.[13] We perform a dynamic analysis using a projection-based approach. We can no longer use the IV approach that we have used in the static analysis, because the instruments are only weakly relevant, especially for longer horizons. Consequently, the first stage coefficients are unreliable. We cannot use these coefficients when computing the impulse response function for longer horizons to study the dynamic effect of a change to government spending. In table 3 below, the second stage coefficients and the F-statistic for instruments are reported for longer horizons.[14] We would expect to see a peak in the F-statistic at a short horizon and then a gradual decrease in the statistic. However, for both instruments this is not the case. The estimated second-stage coefficients are very erratic. The multiplier estimates start positive, then turn negative and turn positive again for longer horizons.

Table 3: Estimated multipliers and F-statistics for longer horizons using IV

	Defense spending		Non-defense spending	
Horizon	Multiplier	F-test	Multiplier	F-test
0	1.166	8.483	2.058	4.042
1	0.977	36.552	-15.824	0.048
2	3.919	6.233	-1.575	7.415
3	-26.400	0.062	-2.372	17.326
4	-9.250	0.157	-20.714	0.243
5	-1.467	2.990	5.441	1.755
6	1.487	0.760	4.995	1.661
7	120.224	0.001	15.715	0.338
8	1.263	3.587	59.537	0.229

Since we cannot use the instruments from the static analysis to estimate impulse responses for longer horizons, we have to rely on a different methodology. We propose to use

[13] An exception is Leduc and Wilson (2013), who calculated longer horizon multipliers for a shock to highway grants.

[14] We use exactly the same regression as in the previous section, where we just shift the dependent variable one period ahead for each horizon.

a projection-based approach. This method makes use of estimated shocks based on forecast errors for the federal government spending categories, i.e. defense contracts or intergovernmental transfers. These estimated government spending shocks should satisfy two conditions. First of all, the shock should be unforecastable using the available information set of economic agents, i.e. all relevant publicly available information. This is likely to be satisfied, because we use an econometric model to construct the forecasts that takes into account as much information as possible. Second, Stock and Watson (2018) note that it is important that the estimated shocks are uncorrelated to (unobserved) past and future shocks, after control variables are included. This condition is formally not testable. However, we include multiple (lagged) control variables and time-fixed effects in the regressions.

4.1 Methodology

We start by estimating the shocks for the different government spending categories: defense contracts and intergovernmental transfers (for non-defense spending). Ideally, we would have used the forecasts for these variables that are available to the government. This is the approach Clemens and Miran (2012) and Leduc and Wilson (2013) followed. However, there are no official forecasts available at the state level for the variables that we use. Therefore, we use an econometric model to obtain forecasts for defense contracts and (categories of) intergovernmental transfers. We estimate a model using a rolling forecast of 10 years to obtain an out-of-sample one-step ahead forecast. Since our data contains a panel of 50 states, there are 500 observations for each forecast. The forecast is state-specific, because we use state-specific values of the right-hand side variables. This forecasting approach is justified because we want to use an ex-ante measure instead of an ex-post measure to create the shock. We therefore use out-of-sample forecasts instead of the residuals from the estimated model.

For the forecast, we use a panel regression with state-level fixed effects. We cannot use time-fixed effects, because this would lead to over-fitting. We include national control variables to control for the national business cycle and aggregate shocks. To obtain the forecast, we use the lags of the dependent variable and K control variables. The dependent variable is either defense contracts for defense spending or a (category of) intergovernmental transfers for non-defense spending. The regression looks as follows:

$$X_{i,t} = \alpha_i + \sum_{\ell=1}^{L} \beta_\ell X_{i,t-\ell} + \sum_{k=1}^{K} \sum_{\ell=1}^{L} \gamma_\ell^k V_{i,t-\ell}^k + \epsilon_{i,t} \tag{4}$$

Here $X_{i,t}$ is defense contracts or intergovernmental transfers. We include state fixed effects α_i, the lags of the dependent variable and K control variables in $V_{i,t}^k$. The following control variables are included: personal income, state government spending, state government tax revenue, federal government spending, the oil price and the real interest rate. All variables have been converted in log real per capita. The lag length is equal to two. Since all variables are trended, we first apply a logarithmic transformation and then regress the variables on a state-specific linear trend. We test whether the results are sensitive to applying a quadratic trend polynomial instead of only a linear trend.

It is important that the forecasts based on this model are qualitatively good forecasts. We check the forecast quality using different indicators for the goodness of fit. These statistics are reported in appendix B. One weakness of this forecast is that it is only based on past information. Hence, it is not possible to filter out all contemporaneous correlation between the government spending variables and business cycle. However, we control for the national business cycle in regressions later. Since state business cycles are synchronized across states, we control for most contemporaneous correlation.

We use the forecasts to compute the shocks for defense contracts and intergovernmental transfers. The shocks are defined as the realized value minus the one-step ahead forecast value. Formally, this can be written as follows:

$$\text{Shock}_{i,t+1} = X_{i,t+1} - \widehat{X}_{i,t+1} \tag{5}$$

The estimated shock represents the unexpected (or unanticipated) change in defense or non-defense spending. This shock causes economic agents to update their expectations about current and future government spending and change their behavior. The approach is similar to the construction of (reduced form) shocks in an SVAR model (Blanchard and Perotti, 2002; Mountford and Uhlig, 2009). The difference is that we do not estimate multiple structural economic shocks jointly, but we only estimate multiple fiscal shocks sequentially. Moreover, we do not directly orthogonalize the estimated shocks to other estimated structural economic shocks. However, we take into account information on other macroeconomic variables when computing the forecast and we control for the national business cycle fluctuations later in the regressions through time-fixed effects.

We use the shocks from the first step to estimate the impact of federal government spending on state-level personal income for longer horizons. To calculate impulse response functions, we use the local projection method by Jordà (2005). The idea of local projections is to estimate a model sequentially at each horizon of the impulse response function, instead of extrapolating a given model into distant horizons, which is common in SVARs. The technique of local projections is more robust to dynamic misspecification and it makes it easier to perform nonlinear estimation.

First, we measure the effect of the shocks on the government spending variable $G_{i,t}$. We estimate the effect of the defense contracts shock on defense contracts and the non-defense (intergovernmental transfers) shocks on state government spending categories. The regressions can be written down in the following way:

$$G_{i,t+h} = \zeta_{i,h} + \phi_{t,h} + \sum_{\ell=1}^{L} \kappa_{\ell,h} G_{i,t-\ell} + \sum_{q=1}^{Q} \sum_{\ell=0}^{L} \xi_{\ell,h}^{q} W_{i,t-\ell}^{q} + \omega_h \text{Shock}_{i,t} + \upsilon_{i,t+h} \tag{6}$$

In this regression at horizon h, the government spending variable $G_{i,t}$ is regressed on state and year fixed effects ($\zeta_{i,h}$ and $\phi_{t,h}$), lags of the dependent variable, Q control variables in $W_{i,t}^{q}$ and the constructed shock in the first step. We include as control variables state-level personal income, the state-level unemployment rate and the state government deficit ratio. The lag length L in the regression is equal to two. As recommended by Stock and Watson (2018), we include lags of the dependent variable to avoid dynamic misspecification of the model and improve estimator efficiency. The control variables and year fixed effects allow to control for contemporaneous correlation with the state business cycle. The regressions are separately estimated for each horizon and each shock. The impulse response function at horizon h for a shock is defined as the estimate for the coefficient ω_h in the regression. We maximize the horizon at eight years. The coefficients for the shock provide an estimate of the impulse response function for defense spending or (categories of) non-defense spending. The confidence interval shows whether the response is significantly different from zero. Alternatively, the significance of the coefficient ω_h can be interpreted as a measure of how well the shock explains the variation in the government spending variable for different horizons, which is comparable to the first stage (weak instrument) F-test in the static approach.

We could have obtained the (asymptotically normal-based) confidence interval directly from the regression. Instead, we use bootstrapped standard errors using the panel moving blocks bootstrap procedure form Gonçalves (2011). The advantage of bootstrapping the coefficients is that this takes into account possible non-normality, heteroskedasticity and autocorrelation. We do not need to know the exact form of the asymptotic distribution. This is especially important for the multiplier estimate later, because the asymptotic distribution for cumulative multipliers is more complicated since it uses the ratio of sums of coefficients.

Unlike the standard block bootstrap that samples across the cross-sectional dimension, the panel moving blocks bootstrap re-samples across the time dimension. This means that in each bootstrap sample all cross-sectional units - in our case states - are included. This approach works even with high cross-sectional correlation, which is likely to be present, because we use states from one country. The bootstrap works as follows. For every iteration, we draw a random vector of starting dates to construct blocks of consecutive years. We set the block length equal to five years. We draw as many blocks as needed to construct a bootstrap sample of the same size as in the original sample. We compute the confidence intervals with the percentile method, using the bias-corrected (centered) bootstrap distribution of the difference between the bootstrapped statistics and the original point estimate. In all cases, we rely on the 95% confidence interval.

The obtained impulse response function and confidence interval can be used to trace the dynamic effect of the shocks on the the government spending variables without modeling explicitly the dynamic interaction between the dependent and independent variables. However, the position of the impulse response function depends on the scaling (or normalization) of the shocks. We scaled the shocks such that the integral of the estimated impulse response function, i.e. the sum of the coefficients, equals one.

Second, we calculate the cumulative (personal income) multiplier. For this, we estimate the following panel regression for state-level personal income $Y_{i,t}$, including year and state fixed effects ($\eta_{i,h}$ and $\theta_{t,h}$) and the same Q control variables $W_{i,t}^{q}$ as in the previous regression:

$$Y_{i,t+h} = \eta_{i,h} + \theta_{t,h} + \nu_{i,h} t + \sum_{\ell=1}^{L} \lambda_{\ell,h} Y_{i,t-\ell} + \sum_{q=1}^{Q} \sum_{\ell=1}^{L} \psi_{\ell,h}^{q} W_{i,t-\ell}^{q} + \rho_h \text{Shock}_{i,t} + \varepsilon_{i,t+h} \tag{7}$$

The shocks for either defense or non-defense spending (categories) obtained in the first step are included. We use two lags of the dependent variable and control variables on the right hand side. The impulse response function shows the dynamic response of personal income to a shock in federal government spending. We can obtain the impulse response directly for each horizon through the estimate of ρ_h. The horizon is again maximized at eight. The confidence interval is obtained using the same moving blocks bootstrap procedure as before.

The two obtained impulse response functions trace the dynamic effect of the shock on government spending and income. We obtain these functions without modeling explicitly the dynamic impulse response between the dependent and independent variables. Moreover, we do not need to model the link between the responses, because it is unnecessary to model the links between the dependent variables endogenously (Jordà, 2005). Unlike in a SVAR model, the impulse responses can be estimated sequentially, because the coefficients in different regressions can be estimated independently.

Finally, we calculate the cumulative multiplier. Our aim is to calculate the cumulative effect on personal income per unit of government spending for different horizons. We calculate the cumulative multiplier by summing the impulse response function of income divided by the sum of the government-spending variable up to horizon h, scaled by the sample average of income $\overline{Y}$ and the government spending variable $\overline{G}$:

$$\text{Cumulative Multiplier}_H = \frac{\sum_{h=1}^{H} \rho_h}{\sum_{h=1}^{H} \omega_h} * \frac{\overline{Y}}{\overline{G}} \tag{8}$$

We bootstrap confidence intervals using the panel moving blocks bootstrap instead of assuming an asymptotic distribution. We sum the coefficients instead of calculating the cumulative effect by accumulating the dependent variable, which is proposed by Chodorow-Reich (2017). Summing the coefficients to obtain the impulse response function is the standard approach in the SVAR literature (Blanchard and Perotti, 2002; Mountford and Uhlig, 2009).

4.2 Results

First, we provide the results for defense spending versus non-defense spending. Our aim is to compare the effect of fiscal shocks of both these spending categories for long horizons. In figure 2 below, we show the estimated responses for defense contracts and state government spending to respectively a defense and non-defense spending shock. The shocks have a strong initial positive effect on the spending variable. Over time, this effect slowly decreases. The shocks are scaled such that the sum of the impulse response equals one. The shocks do not have a permanent effect on government spending, but contain a one-time fiscal impulse. The response for defense federal spending shows a stronger initial effect, whereas the response for non-defense spending appears to be more persistent in the first years. The responses show that there is a significant positive effect on both defense contracts and state spending for 4 years. The figure further shows the response of personal income to a defense or non-defense shock. We find that a defense shock has an accumulating positive effect on income. Instead, the effect of non-defense spending is significantly positive in short horizons, but negative in longer horizons. Hence, we find a very different response of income to both shocks.

Figure 2 furthermore shows estimates of the cumulative multiplier and the 95% bootstrapped confidence intervals as shaded areas. We find that the multiplier for a non-defense shock starts above 1 and continues increasing until 2.5 after 8 years. The multiplier for a

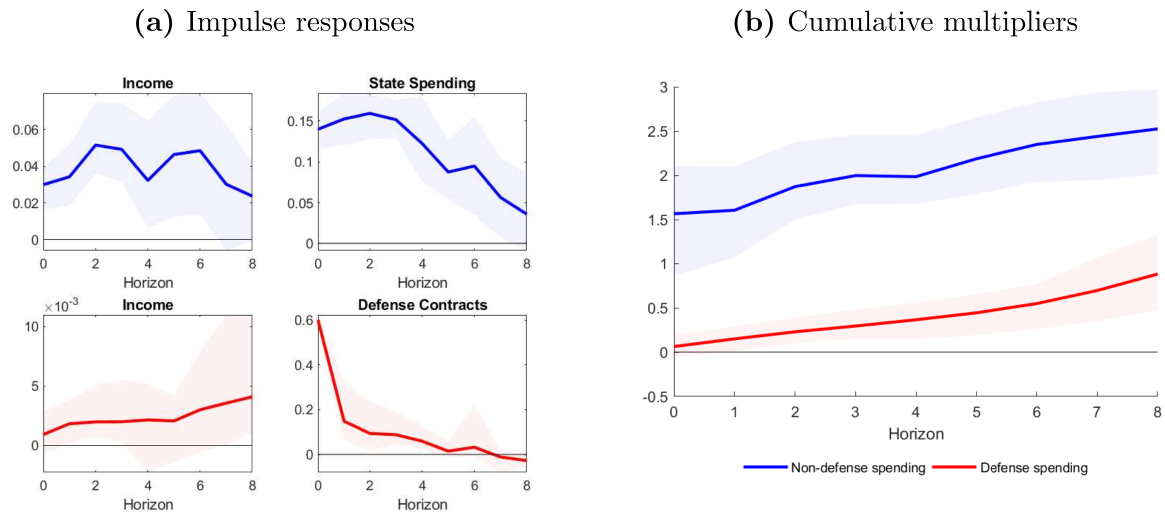

Figure 2: Estimates and bootstrapped 95% confidence intervals for the dynamic analysis, using a state-specific linear trend and state and time-fixed effects. The non-defense spending shock is shown in blue and the defense spending shock is shown in red.

defense shock starts around zero and increases towards 1 after 8 years. We find a significant difference between the effect of the two shocks. The cumulative multiplier for a non-defense shock is higher than for a defense shock. We later also check which category of non-defense spending drives this higher multiplier estimate. Compared to the static analysis, it now takes eight years for the multiplier of a non-defense shock to reach a value of 1.0. In the static analysis we observed a multiplier above 1.0 already after two years. Some other findings are similar, such as the higher multiplier estimates for a non-defense shock. The dynamic analysis now furthermore shows that there is a significant difference between the shocks.

We also check whether the shocks affect the intended spending channel. In figure 3, we show the effect of the non-defense shock to defense spending and the defense shock to non-defense spending. The non-defense shock does not affect defense spending significantly. However, there is a significant increase in non-defense spending to a defense shock after 2 years. The explanation for this result is that the economic stimulus of defense spending affects state spending positively, because state spending is largely pro-cyclical.

To check the robustness of these results, we check whether the results are robust to including a quadratic trend, because series show concavity over time. We detrend the variables

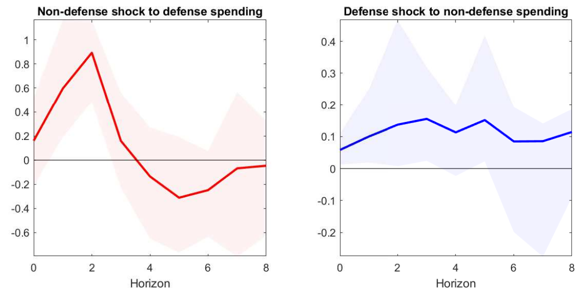

Figure 3: Impulse responses for the shocks on the other spending channel, i.e. effect of the non-defense shock on defense spending and the defense shock on non-defense spending. Estimates and bootstrapped 95% confidence intervals are computed using a state-specific linear trend and state and time fixed effects.

using a quadratic trend polynomial at the state level (see appendix C). The magnitude of the response of income to a non-defense shock decreases. Furthermore, both multipliers decrease and the multiplier for a non-defense shock becomes more stable over time and stays around 1.0. The effect of a non-defense shock becomes weaker. The conclusion remains that an non-defense spending shock has a larger positive effect on personal income than a defense spending shock.

Second, we investigate the effect of different categories of non-defense spending. Figure 4 provides estimation results for different categories of non-defense spending: education, health, highways and Medicaid. We estimate the impulse response functions for different categories of state government spending. We find that all categories respond significantly positive to a shock in these spending categories. We also calculate the cumulative multiplier for the different categories of non-defense spending. We find especially for education and Medicaid a significantly positive multiplier. The multiplier of highways is only significant for longer horizons.

We checked whether the results are robust to including a quadratic trend, because the series show concavity over time (see appendix C). The impulse responses are unaffected by including a quadratic trend. However, the confidence intervals become much wider. We do not find strong evidence for a significant positive multiplier for one of the non-defense spending categories.

(a) Impulse responses (b) Cumulative multipliers

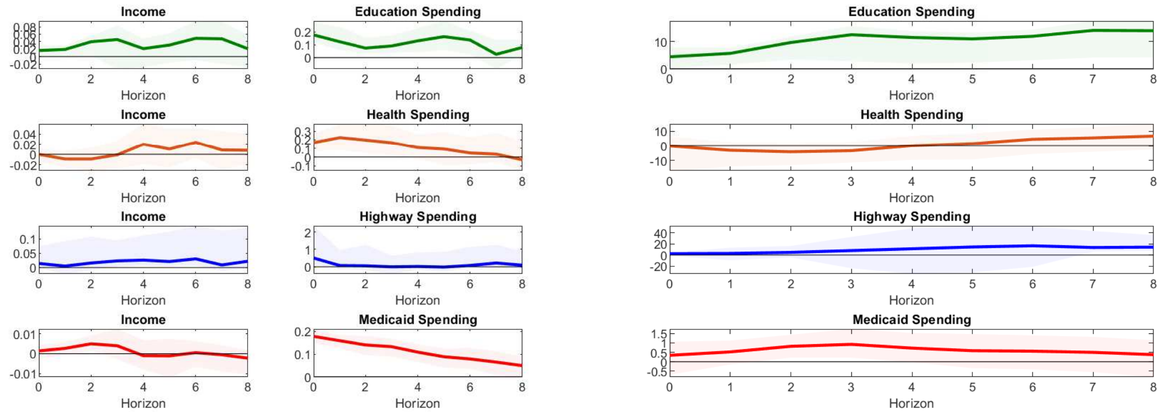

Figure 4: Estimates and bootstrapped 95% confidence intervals for a shock to different categories of grants: Education grants, Health care grants, Highway grants or Medicaid grants, using a state-specific linear trend and state and time-fixed effects.

5 Conclusion

Many studies have provided estimates for the local fiscal multiplier for federal government spending in a fiscal union. However, no study has compared the effectiveness of different categories of federal government spending. This paper provides comparable estimates for different categories of spending. By using U.S. data on defense contracts and federal intergovernmental transfers to state governments, we estimate the local fiscal multiplier for different types of defense and non-defense spending for the entire post-war period. In the static IV analysis, we estimate a two-year (short-horizon) impact multiplier around 1.2 for defense spending and around 2.1 for non-defense spending. We use a projections-based approach for the dynamic analysis, because the instruments are only weakly relevant for longer horizons. The estimated multipliers for non-defense spending are higher than for defense spending. We also estimate the multipliers for categories of non-defense spending, but we do not find clear evidence that a particular category of spending drives the result.

This paper adds empirical evidence to the growing body of literature on the local fiscal multiplier. Its main contribution is the new strategy to capture indirect federal government spending in a broad sense with intergovernmental grants. Our results indicate that non-defense spending by the federal government has a stronger positive effect on personal income. We find this result both in the static and dynamic analysis. However, it remains difficult to estimate multipliers for the separate categories of non-defense spending. Spending for these categories is probably too persistent. This makes it hard to estimate shocks that affect these spending categories on impact. Moreover, non-defense spending by the government is intended to promote long-term welfare instead of short-run gains. It remains puzzling why we find significant positive multipliers for total non-defense spending, but not for the individual categories.

Future research could determine what drives the different multipliers between the spending categories. Since our empirical analysis focuses exclusively on the U.S., we cannot claim with certainty that the results carry over to other fiscal unions. We use the specific structure of government spending in the U.S. in our methodology. Hence, we cannot claim that the differences between the multipliers for some categories are unaffected by the specific channel of government spending, i.e. direct federal spending or indirect federal spending. Our method does not allow to disentangle the channel and category of government spending. However, with the methodology proposed, future research can perform this analysis for other countries. We focus only on the effect of federal government spending on personal income, but there are many other on, for example, the labour market or the financial market. Knowing more about these spill-overs could also help to explain why we estimate different multipliers for defense and non-defense spending.

References

Adelino, M., Cunha, I., and Ferreira, M. (2017). The economic effects of public financing: Evidence from municipal bond ratings. *The Review of Financial Studies*, 20 (9):3223 – 3268.

Auerbach, A. and Gorodnichenko, J. (2012). Measuring the output responses to fiscal policy. *American Economic Journal: Economic Policy*, 4 (2):1 – 27.

Barro, R. and Redlick, C. (2011). Macroeconomic effects from government purchases and taxes. *Quarterly Journal of Economics*, 126:51 – 102.

Bartik, T. (1991). *Who Benefits from State and Local Economic Development Policies?*. W.E. Upjohn Institute for Employment Research, Michigan.

Biolsi, C. (2015). Local effects of a military spending shock: Evidence from shipbuilding in the 1930s. *Working Paper, University of Houston.*

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 (4):1329–1368.

Chodorow-Reich, G. (2017). Geographic cross-sectional fiscal spending multipliers, what have we learned? *NBER Working Paper*, 23577.

Chodorow-Reich, G., Feiveson, L., Liscow, Z., and Woolston, W. (2012). Does state fiscal relief during recessions increase employment? *American Economic Journal: Economic Policy*, 4 (3):118–145.

Clemens, J. and Miran, S. (2012). Fiscal policy multipliers on subnational government spending. *American Economic Journal: Economic Policy*, 4 (2):46–68.

Conley, T. and Dupor, B. (2013). The american recovery and reinvestment act: Soleley a government job program? *Journal of Monetary Economics*, 60:535–549.

Dupor, B. and Guerrero, R. (2017). Local and aggregate fiscal policy multipliers. *Journal of Monetary Economics*, 92:16–30.

FFIS (2016). Summary of state matching and MOE requirements. *Special Analysis 16-03, Federal Funds Information for States, Washington DC.*

FFIS (2018). Grants 101: An introduction to federal grants for state and local governments. *Primer Update, Federal Funds Information for States, Washington DC.*

Fishback, P. and Kachanovskaya, V. (2015). The multiplier for federal spending in the states during the Great Depression. *The Journal of Economic History*, 75 (1):125–162.

Gamkar, S. and Oates, W. (1996). Asymmetries in the responses ot increases and decreases in intergovernmental grants: Some empirical findings. *National Tax Journal*, 49 (4):501–512.

Gonçalves, S. (2011). The moving blocks bootstrap for panel linear regression models with individual fixed effects. *Econometric Theory*, 27 (5):1048–1082.

Gordon, N. (2004). Do federal grants boost school spending? Evidence from Title I. *Journal of Public Economics*, 88:1771–1792.

Hausman, J. (2016). Fiscal policy and economic recovery: The case of the 1936 Veterans Bonus. *American Economic Review*, 106 (4):1100–1143.

Ilzetzki, E., Mendoza, E., and Vgh, C. (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.

Klarner, C. (2013). State partisan balance data, 1937 - 2011. *Harvard University.*

Knight, B. (2002). Endogenous federal grants and crowd-out of state-government spending: Theory and evidence from the Federal Highway Aid Program. *American Economic Review*, 92 (1):71–92.

Leduc, S. and Wilson, D. (2013). Roads to prosperity or bridges to nowhere? Theory and evidence on the impact of public infrastructure investment. *NBER Macroeconomics Annual 2012*, pages 89–142.

Mountford, A. and Uhlig, H. (2009). What are the effects of fiscal policy shocks? *Journal of Applied Econometrics*, 24 (6):960–992.

Nakamura, E. and Steinsson, J. (2014). Fiscal stimulus in a monetary union: Evidence from US regions. *American Economic Review*, 104 (3):753–792.

Nesbit, T. and Kreft, S. (2009). Federal grants, earmarked revenues, and budget crowd-out: State highway funding. *Public Budgeting & Finance*, 29 (2):94–110.

Oates, W. (1999). An essay on fiscal federalism. *Journal of Economic Literature*, 37:1120–1149.

Ramey, V. (2011). Identifying government spending shocks: It's all in the timing. *Quarterly Journal of Economics*, 126 (1):1–50.

Ramey, V. and 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.

Shoag, D. (2016). The impact of government spending shocks: Evidence on the multiplier from state pension plan returns. *Working paper, Harvard University*.

Stock, J. and Watson, M. (2018). Identification and estimation of dynamic causal effects in macroeconomics using external instruments. *The Economic Journal*, 128:917–948.

Suárez Serrato, J. and Wingender, P. (2016). Estimating local fiscal multipliers. *NBER Working Paper*, No. 22425.

Wilson, D. (2012). Fiscal spending jobs multipliers: Evidence from the 2009 American Recovery and Reinvestment Act. *American Economic Journal: Economic Policy*, 4 (3):251–282.

Appendix A Robustness checks static analysis

A.1 2-year growth rate instruments

Table 4: Static results for defense versus non-defense spending

	(1) Pers. Income	(2) Pers. Income	(3) Pers. Income	(4) Pers. Income
Defense Spending	0.9489*** (0.2454)	1.0165*** (0.2551)	-	-
Non-defense Spending	-	-	2.3642 (2.3654)	2.3059 (2.5461)
Fixed effects	State, Year	State, Year	State, Year	State, Year
Political controls	No	Yes	No	Yes
First stage F-stat.	37.66	35.95	1.94	1.61
N	2450	2401	2450	2401

Robust SEs in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 5: Static results for categories of non-defense spending

	(1) Pers. Income	(2) Pers. Income	(3) Pers. Income	(4) Pers. Income
Education Spending	4.8023 (3.9922)	-	-	-
Highway Spending	-	-3.4734* (1.8095)	-	-
Health Spending	-	-	125.2987 (221.5589)	-
Medicaid Spending	-	-	-	6.1650** (3.0497)
Fixed effects	State, Year	State, Year	State, Year	State, Year
Political controls	No	No	No	No
First stage F-stat.	0.98	25.90	0.39	10.82
N	2450	2450	2450	2450

Robust SEs in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

A.2 Political controls for non-defense spending

Table 6: Static results for categories of non-defense spending

	(1) Pers. Income	(2) Pers. Income	(3) Pers. Income	(4) Pers. Income
Education Spending	17.2212	-	-	-
	(91.6772)			
Highway Spending	-	-3.8620	-	-
		(2.6758)		
Health Spending	-	-	145.8287	-
			(353.6192)	
Medicaid Spending	-	-	-	9.0944**
				(3.8845)
Fixed effects	State, Year	State, Year	State, Year	State, Year
Political controls	Yes	Yes	Yes	Yes
First stage F-stat.	0.03	9.71	0.22	9.15
N	2401	2401	2401	2401

Robust SEs in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Appendix B Forecasting model for dynamic analysis

We construct a forecast for the federal spending variables defense contracts and (the categories of) intergovernmental transfers for non-defense spending. This forecast is based on a 10-year rolling-window panel regression with fixed effects using two lags of the following variables: personal income, state-government spending, state-government tax revenue, federal-government spending, the oil price and the real interest rate. Based on this regression, we compute a one-step ahead out-of-sample forecast $\hat{X}_{i,t}$. This forecast is different from the residual of the model. Hence, the model can be very good at making in-sample predictions, but perform bad when making forecasts. The forecast quality of the model can be evaluated using the root mean squared forecast error (RMSE):

$$RMSE = \sqrt{\frac{1}{N}\sum_{n=1}^{N}(X_{i,t} - \hat{X}_{i,t})^2} \tag{9}$$

Here N is the total number of observations in the sample across all states. The forecast is made after detrending the variables at the state level. We check the quality of the forecast both when using a linear and quadratic trend polynomial. In table 7 below, we report the RMSE of both the linear and the quadratic forecast. We compare this to the naive forecast, where we use last year's value as a forecast.

Table 7: Root mean squared forecast error

Spending category	RMSE Naive	RMSE Linear	RMSE Quadratic
Defense contracts	0.41608579	0.41679533	0.44428328
IGT	0.09293558	0.14423047	0.15460218
IGT Education	0.16350102	0.14629919	0.15483091
IGT Health	0.23654438	0.22437945	0.22832302
IGT Highways	0.45948542	0.55787661	0.56044479
IGT Medicaid	0.12127707	0.16891916	0.16859381

The results in the table show that including a quadratic trend does not improve the forecast quality. The model performs better in forecasting IGT than defense contracts and in particular education and Medicaid grants. However, for some spending categories, the naive forecast performs even better. Especially for total IGT, Highways and Medicaid, the RMSE naive forecast is lower. Additionally, we check the role of forecast window. In the standard specification, we use a 10-year rolling window. Below in table 8, we show the results when using a 15-year window. A longer window improves the forecast.

Table 8: Root mean squared forecast error using a 15-year window

Spending category	RMSE Naive	RMSE Linear	RMSE Quadratic
Defense contracts	0.41608579	0.37690784	0.36993414
IGT	0.09293558	0.08870693	0.08243247
IGT Education	0.16350102	0.11989164	0.11867427
IGT Health	0.23654438	0.21712125	0.20157065
IGT Highways	0.45948542	0.52693242	0.52098474
IGT Medicaid	0.12127707	0.11854984	0.11382192

Indeed, we find that the forecast improves a lot when using a longer window. All forecasts improve and in all cases except Highways the linear forecast now beats the naive forecast. Moreover, in all cases the quadratic forecast has a lower RMSE than the linear forecast.

Appendix C Dynamic analysis with quadratic trend

Defense vs. non-defense spending

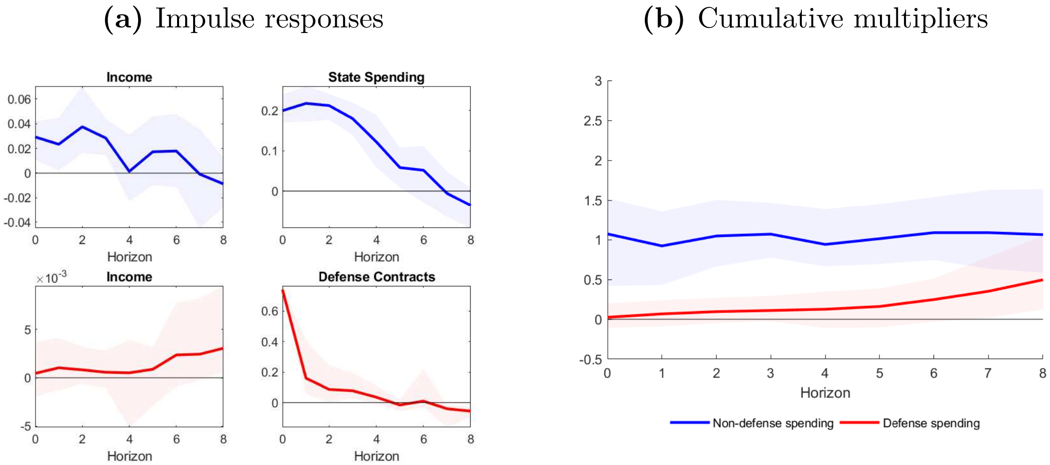

Figure 5: Estimates and bootstrapped 95% confidence intervals for the dynamic analysis, using a state-specific quadratic trend and state and time-fixed effects. The non-defense shock is shown in blue and the defense shock is shown in red.

Categories of non-defense spending

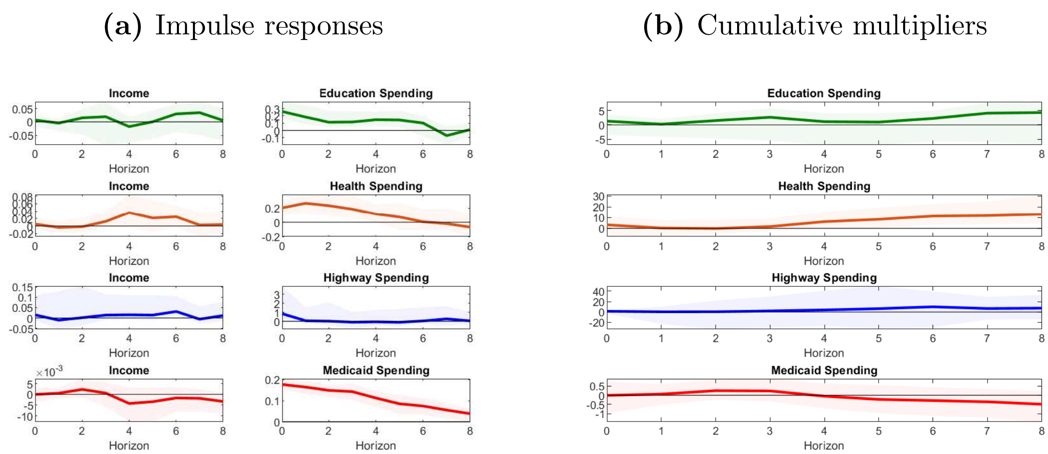

Figure 6: Estimates and bootstrapped 95% confidence intervals for a shock to different categories of grants: Education grants, Health care grants, Highway grants or Medicaid grants, using a state-specific quadratic trend and state and time-fixed effects.