# Estimating The Missing Intercept

Christian Matthes, Naoya Nagasaka, and Felipe Schwartzman*

October 16, 2024

**Abstract**

Microeconometric approaches to answer macroeconomic questions regularly use time fixed effects. This leads to the well-known 'missing intercept' problem because fixed effects soak up average aggregate effects. As such, these results cannot be used to directly address policy questions requiring knowledge of policies' aggregate effects. We present a statistical approach that leverages knowledge of these microeconometric results to jointly identify aggregate and idiosyncratic effects of changes in policy. We then apply our methodology to study government spending multipliers (Nakamura and Steinsson, 2014).

JEL Classification: C11, C50, E62, H50, R12

Keywords: Fixed Effects, Aggregate Effects, Government Spending, Regional Data, Bayesian Analysis

*Indiana University (Matthes & Nagasaka) & Federal Reserve Bank of Richmond. Contact information: matthesc@iu.edu, naonagas@iu.edu, felipe.schwartzman@rich.frb.org. We would like to thank Mark Watson as well as seminar participants in Berlin, Kiel, München, DNB, ECB, ESM, Tübingen, and the NY Fed for very helpful feedback. The views expressed in this paper are those of the authors and do not necessarily reflect those of the Federal Reserve Bank of Richmond or the Federal Reserve System.

# 1 Introduction

Modern macroeconomic research relies increasingly on panel datasets featuring variation across regions, households, or firms. This idiosyncratic variation allows researchers to leverage microeconometric tools designed to credibly identify the effects of policies. A key part of this toolkit is time fixed effects. These are helpful for identification because they remove the effect of aggregate shocks that affect all variables simultaneously.

However, as is well recognized, time fixed effects also remove the aggregate response to policy changes. As a result, these methods can only uncover the idiosyncratic or local effects of policy changes *net* of those aggregate effects. Those tools are therefore inadequate to *directly* answer questions about aggregate effects of policies. A typical response in the literature is to use the estimates obtained with fixed effects to calibrate fully specified dynamic equilibrium models (see, for example, Nakamura and Steinsson (2014)).[1] While informative, those strategies provide estimates of aggregate effects that depend on the specifics of the structural model.

We propose a methodology to estimate the aggregate effects of policies by merging the microeconometric tools that provide sharp identification at the idiosyncratic level with time-series methods that estimate aggregate effects without the strong cross-equation restrictions of dynamic equilibrium models. In particular, we show how to incorporate estimates of local effects net of aggregate effects into a time series model that jointly describes the evolution of aggregate and local economic outcomes. This allows us to simultaneously use variation in both the cross-section and time series dimensions to sharpen estimates. Furthermore, our approach allows researchers to combine the microeconometric approach with identification assumptions traditionally used in the time series literature, such as zero, sign and magnitude restrictions (Christiano et al., 1999; Uhlig, 2005; Canova and Nicolo, 2002; Faust, 1998; Amir-Ahmadi and Drautzburg, 2021) on the impact of shocks, as well as instruments for aggregate shocks (Mertens and Ravn, 2013; Plagborg-Møller and Wolf, 2021).

Intuitively, the method identifies aggregate shocks from the behavior of idiosyncratic units by exploiting as input the effects estimated using microeconometric tools. For example, suppose that a national government spending shock increases spending by more in Wisconsin than in New York, and implies an increase in output in Wisconsin that is larger than in New York by an amount estimated using microeconometric methods. We can then infer the trajectory of that shock from the relative movements of government spending and output in those two regions. This estimated aggregate shock can then be used to infer the aggregate

[1] An alternative pursued by Chodorow-Reich (2019, 2020) is to derive situations in which the estimated local equilibrium effects can be interpreted as bounds on the aggregate or general equilibrium effects.

output multiplier. In practice, we perform the estimation within a single Bayesian model. This allows us to (i) not dogmatically impose results from the applied micro literature but use them to inform priors, (ii) perform the estimation simultaneously, not sequentially, thus making the best and consistent use of all information available, and (iii) to use priors for the purpose of regularization (for example, the use of a Minnesota-type prior for coefficients in our time series model (Doan et al., 1984)).

We demonstrate the methodology in an application. We revisit the famous Nakamura and Steinsson (2014) analysis of fiscal multipliers across U.S. states. We use this application to study the impact of various modeling choices we have made in the benchmark scenario and provide evidence that the aggregate government spending multiplier is less than unity with high probability. We disentangle where this result comes from, and while it is generally robust, we do find that prior information is needed to arrive at meaningfully tight posterior bands. We provide both statistical and economic foundations for this result. In particular, it echoes results in Nakamura and Steinsson (2014), where the authors find that many different structural models that imply vastly different aggregate multipliers are consistent with the same local effects of aggregate feedback.

Although the missing intercept problem is distinct from other econometric issues related to cross-sectional multipliers discussed by Canova (2022), our approach is general enough to not fall victim to the issues discussed in that paper (i.e., we allow for heterogeneity across cross-sectional units). Our paper is related but distinct from previous work on the missing intercept problem in Wolf (2023), which provides results under which these micro-based local effects can be added to a macro/time-series-based estimate of the aggregate effect to arrive at the total effect at the local level. We instead leverage micro-based estimates to jointly estimate aggregate and total local effects.

Sarto (2024) also leverages regional data to uncover aggregate effects, exploiting, as we do, a factor structure in the data, but then combines this factor structure with exclusion restrictions to achieve identification. Our approach is complementary in that it directly leverages microeconometric estimates,and connects those results to the large literature on identification in Vector Autoregressions (VARs) and the time-series literature more generally, identification approaches from which can also be easily incorporated into our approach. Furthermore, since we use a Bayesian approach, we have a natural avenue to introduce regularization via priors, which can be helpful in high-dimensional parameter spaces such as in our applications. Finally, our Bayesian approach allows us to dogmatically impose exclusion restrictions along the lines of Sarto (2024) or use them to center a non-degenerate prior. Interestingly, Sarto (2024) finds an aggregate government spending multiplier that is broadly in line with our findings. The idea of exploiting variation at various levels of aggregation to identify effects at the aggregate level has recently become more popular - Gabaix and Koijen (2023) show how to exploit variability in large cross-sectional units to derive instrumental variables. A related full information approach is developed in Baumeister and Hamilton (2023). Our approach can exploit information found in Bartik instruments and, as such, builds on the growing literature studying these instruments (Bartik, 1991; Goldsmith-Pinkham et al., 2020; Borusyak et al., 2021).

The remainder of the article is structured as follows: Section 2 highlights how microeconometric studies deliver estimates of local effects and how they can be linked to aggregate effects. Section 3 discusses our time series model that can leverage such estimates of local effects for identification. Section 4 discusses the link between microeconometric identification and identification in our approach in more detail. Section 5 provides an application of our approach, building on Nakamura and Steinsson (2014). Section 6 provides Monte Carlo evidence on the performance of our approach, and Section 7 concludes.

# 2 The Missing Intercept: An Example

To set the stage, we now build an economy inspired by Moll (2021), designed to discuss the missing intercept problem. Using this example, we describe what objects are identified from cross-sectional variation and how we can exploit this information to estimate both aggregate effects and total effects for each cross-sectional unit. The example is purposefully simplified for ease of exposition, with many of the special assumptions relaxed in the full model described in Section 3.

For concreteness, the example revolves around the use of local data to estimate the effects of government spending on output, though this is merely a matter of labeling variables. The example can be easily recast for the problem of estimating the effect of wealth shocks on employment or several other problems of interest to macroeconomists.

We consider a set of $N$ regions (which in other examples could be sectors, households, firms etc). In each region $i \in \{1, ..., N\}$, the output is determined by government spending in the region $i$, $g_{it}$, as well as aggregate government spending $G_t$, regional-specific exogenous factors $\varepsilon_{it}^{y}$ and other aggregate shocks $\eta_t^Y$ such as shocks to Total Factor Productivity or monetary policy.

$$y_{it} = \gamma g_{it} + \theta G_t + \eta_t^Y + \varepsilon_{it}^{y}, \tag{1}$$

where aggregate government spending is itself a function of shocks summarized in $\eta_t^Y$ and an aggregate government spending shock $\eta_t^G$:

$$G_t = \eta_t^G + \rho \eta_t^Y$$

While local output depends on local government expenditures through the usual government demand multiplier channels, captured in $\gamma$, it may also depend on aggregate spending $G_t$ either because there is an endogenous monetary response, through trade between regions, or other country-wide general equilibrium channels with those captured in $\theta$.

Panel studies often focus on estimating $\gamma$. Estimates of $\gamma$ may be of interest for several reasons. First, they may be useful in their own right for the evaluation of local policies. Second, they can provide discipline to identify deep parameters of interest, as in Nakamura and Steinsson (2014). Third, in combination with separate estimates of $\theta$, they provide the local impact of aggregate fiscal shocks, as in Wolf (2023). Finally, if one is able to a priori put a sign to $\theta$, they provide useful bounds on the total effects of those aggregate fiscal shocks, as proposed by Chodorow-Reich (2019).

The main identification challenge for panel studies is that, either because of omitted variables or reverse causation, $g_{it}$ may be determined by the same idiosyncratic and aggregate factors that determine $y_{it}$, namely, for all $i \in \{1, ..., N\}$

$$g_{it} = \beta_i G_t + \xi \varepsilon_{it}^y + \varepsilon_{it}^g. \tag{2}$$

where now $\varepsilon_{it}^g$ are idiosyncratic government spending shocks affecting local government spending, and $\beta_i$ and $\xi$ are coefficients.

We define aggregate variables as averages of the idiosyncratic variables, as is appropriate in most applications where variables are expressed in growth rates, logarithms, or per-capita terms. We assume, following Moll (2021), that local shocks do not have a direct effect on aggregate outcomes: $\frac{1}{N}\sum_{i=1}^N \varepsilon_{it}^g = \frac{1}{N}\sum_{i=1}^N \varepsilon_{it}^y = 0$. As a result, we can write aggregate variables as

$$\frac{1}{N}\sum_{i=1}^N g_{it} = G_t = \eta_t^G + \rho \eta_t^Y \tag{3}$$

$$\frac{1}{N}\sum_{i=1}^N y_{it} = Y_t = (\gamma + \theta)\, G_t + \eta_t^Y \tag{4}$$

where we use the fact that the aggregation requires $\frac{1}{N}\sum_i \beta_i = 1$. This implies that a time-fixed effect for local output and government spending recovers the aggregate version of those variables.

A common strategy in panel studies is as follows:

1. Use time-effects to control for the effects of $\eta_t^y$, the common shocks to $g_{it}$ and $y_{it}$, obtaining the system of equations

$$y_{it} - Y_t = \gamma(g_{it} - G_t) + \varepsilon_{it}^y \tag{5}$$
$$g_{it} - G_t = (\beta_i - 1)G_t + \xi\varepsilon_{it}^y + \varepsilon_{it}^g \tag{6}$$

2. Obtain a measure of $\beta_i - 1$. For example, Nakamura and Steinsson (2014) use either the coefficients of an OLS regression of $g_{it} - G_t$ on $G_t$ or shares of military expenditures in local GDP.

3. Use $(\beta_i - 1)G_t$ as an instrument for $g_{it} - G_t$ in equation (5)

This strategy successfully identifies $\gamma$ under the assumption that $\beta_i$ is uncorrelated with $\varepsilon_{it}^y$.

The missing intercept problem occurs because macroeconomic policy analysis requires understanding the *total* effect of government spending on output, $\gamma + \theta$. However, in the methodology above, the term $\theta$ is absorbed by time effects and cannot be estimated.

To estimate $\theta$ one needs to bring attention back to the aggregate equations (3) and (4). However, inference is complicated by the fact that both $G_t$ and $Y_t$ are functions of the same set of macroeconomic shocks, summarized in $\eta_t^Y$. Prior literature on fiscal policy has focused on imposing additional identification assumptions on aggregate data. A particularly fruitful strategy has been to find instruments that plausibly generate variations in $G_t$ directly and on $Y_t$ only indirectly, through $G_t$. For example, Ramey (2011); Auerbach and Gorodnichenko (2012) use instruments such as military spending for the case of government spending multiplier.

Our main contribution is to propose an alternative strategy that does not require identification assumptions beyond what is used to estimate $\gamma$ in the panel approach above. To understand our strategy, it is useful to rewrite equations in terms of shocks only, with the equations for deviations of $g$ and $y$ from aggregates, (6) and (5), rewritten as:

$$g_{it} - G_t = B_{GG}^i \eta_t^G + B_{GY}^i \eta_t^Y + \xi\varepsilon_{it}^y + \varepsilon_{it}^g, \tag{7}$$
$$y_{it} - Y_t = B_{YG}^i \eta_t^G + B_{YY}^i \eta_t^Y + (1 + \gamma\xi)\varepsilon_{it}^y + \gamma\varepsilon_{it}^g \tag{8}$$

where $B_{GG}^i = (\beta_i - 1)\lambda$ and $B_{GY}^i = (\beta_i - 1)\rho$, $B_{YG}^i = \gamma B_{GG}^i$ and $B_{YY}^i = \gamma B_{GY}^i$.

From these equations, it is apparent that $g_{it} - G_t$ and $y_{it} - Y_t$ have a factor structure. That is, they are functions of aggregate shocks $\eta_t^G$ and $\eta_t^Y$ that affect values in each region with different loadings ($B_{GG}^i, B_{GY}^Y$ etc) and idiosyncratic shocks $\varepsilon_{it}^y$ and $\varepsilon_{it}^g$ which are region-specific. Econometric theory makes clear that one can identify the space spanned by the aggregate factors $\eta_t^G$ and $\eta_t^Y$ with very little information beyond the time-series for the observed variables and the number of factors.

The key identification challenge is disentangling $\eta_t^G$ from $\eta_t^Y$. Fortunately, this can be done given information on $B_{GG}^i$ and $B_{YG}^i$. Those are functions of $\beta_i$ used to construct the instrument in the panel estimation described above and of the partial effect $\gamma$, obtained in that same estimation. As shown in Section 4, that information is sufficient for identification of $\eta_t^G$.[2] The estimate of $\eta_t^G$ can then be used to obtain an estimate of $\gamma + \theta$ from aggregate data.

The model described above is simplified for expositional purposes. It assumes that $\eta_t^Y$ and $\eta_t^G$ can be differenced out from equations (1) and (2), which may be generally not true since responses to aggregate shocks and the general equilibrium feedback through $\theta G_t$ may be heterogeneous. This complicates the estimation of $\gamma$ in microeconometric studies, perhaps requiring the introduction of additional controls or other corrections (as emphasized by Canova (2022)). However, they do not affect our ability to identify $\eta_t^G$ given those estimates. The model also does not allow for dynamics, which can matter for estimates of the government spending multiplier or other macroeconomic effects. Lastly, one may be concerned about error in the $\beta^i$ and $\gamma$'s obtained from microeconometric studies, or that the $\gamma$'s are heterogeneous across units.

In the next section, we develop a more flexible framework featuring a richer version of local (1, 2) and aggregate (3, 4) equations that allow for those complications. In particular, the model incorporates persistence, heterogeneous responses to other aggregate shocks across cross-sectional units, imperfect knowledge of $B_{GG}^i$ and $B_{YG}^i$.[3] It also safeguards against misspecification or estimation error in $B_{GG}^i$ and $B_{GY}^i$ by using this information to establish a prior that we use for Bayesian inference, rather than imposing those dogmatically on either the aggregate or idiosyncratic effects of this shock.

The objects of interest in our time series model are the impulse responses to an aggregate government spending shock $\eta_t^G$ and the ratio of those impulse responses for output and aggregate government spending – the fiscal multiplier. In the example in this section, the fiscal multiplier equals $\gamma + \theta$, which was the focus of this discussion.

[2]The required assumption is that there is indeed heterogeneity in $\lambda_i^G$, similar to assumptions in Wolf (2023) and Sarto (2024)

[3]In practice, our approach jointly estimates parameters for the entire system at once.

# 3 Time Series Model

We now describe the full time-series model, which generalizes the model in Matthes and Schwartzman (2023).

The model provides a flexible data-generating process that jointly describes micro- and macroeconomic dynamics. It consists of a block for aggregate data and blocks for idiosyncratic units such as localities, sectors, etc. In both levels of aggregation, we use variants of Vector Autoregressive (VAR) models. The blocks are linked via aggregate variables and structural shocks, allowing for rich patterns of comovement while remaining parsimonious in terms of parametrization.

For each of $I$ idiosyncratic units, we track $K$ unit-specific variables such as output, expenditures, or prices. Those can be combined into an equal number of aggregate variables, to which we can add aggregate-only variables such as policy interest rates, national government spending, or stock price indices, for a total of $N > K$ aggregate variables.

The model explains those variables in terms of $R$ aggregate shocks with $R << I$ as well as shocks specific to each aggregate or idiosyncratic variables. We now describe the aggregate and idiosyncratic blocks in detail.

**Block 1: Aggregate**

The aggregate block can be written, in vector form, as

$$X_t^{agg} = \mu^{agg} + \sum_{l=1}^{L} A_l^{agg} X_{t-l}^{agg} + B^{agg}\eta_t + \varepsilon_t, \tag{9}$$

where $X_t^{agg}$ is an N dimensional vector collecting observed aggregate endogenous variables, $\eta_t \sim N(0, I)$ is a $R$ dimensional vector of unobserved aggregate shocks with entries (where we allow for $N \geq R$), and $\varepsilon_t \sim N(0, \Sigma^{agg})$ collects other shocks affecting aggregate variables as well as measurement error. The aggregate block features $L$ lags. $\mu^{agg}$, $A_l^{agg}$ and $B^{agg}$ are conformable vectors and matrices of parameters to be estimated. $B^{agg}$ captures effects of structural shocks on aggregate variables on impact. We generally denote entries in $B^{agg}$ by $B_{nr}^{agg}$, where $n \in \{1, ..., N\}$ indexes the variable and $r \in \{1, ..., R\}$ the aggregate shock unless noted otherwise.

**Block 2: Idiosyncratic**

For each idiosyncratic unit $i$, the idiosyncratic block can be written, in vector form, as

$$X_t^i - X_t^{agg} = \mu^i + \sum_{l=1}^{L^{agg}} A_l^i X_{t-l}^{agg} + \sum_{l=1}^{L^{reg}} C_l^i X_{t-l}^i + B^i \eta_t + \varepsilon_t^i, \quad i = 1, \cdots, I \tag{10}$$

where $X_t^i$ is an K-dimensional vector including the idiosyncratic endogenous variables, and $\varepsilon_t^i \sim N(0, \Sigma^i)$ is assumed to be independent across idiosyncratic units and independent of any shock at the aggregate level, though not necessarily across variables within idiosyncratic units. $L^{agg}$ and $L^{reg}$ denote the number of lags of aggregate and idiosyncratic variables. $\mu^i$, $A_l^i$, $C_l^i$ and $B^i$ are conformable vectors and matrices of parameters. $B^i$ captures the effects of aggregate shocks *net of the effect on aggregate variables* which also has the interpretation of effects obtained after controlling for *time fixed effects*. We denote entries in $B^i$ by $B_{kr}^i$, where $k \in \{1, ..., K\}$ indexes the variable and $r \in \{1, ..., R\}$ the aggregate shock.

$X_t^{agg}$ is subtracted from the left hand side in order to account for the time fixed effect. While we assume here for simplicity that the variables in $X_t^{agg}$ are the direct aggregate counterpart of the local variables in $X_t^i$, we can easily accommodate more aggregate variables.[a] Spillovers across regions occur due to aggregate shocks $\eta_t$ or contemporaneous and lagged aggregate variables $X_t^{agg}$.[b]

[a]In that case we simply need to modify the left-hand side of Equation (10) to be $X_t^i - SX_t^{agg}$, where $S$ is a selection matrix that selects those observables that we can measure both at the aggregate and local levels.

[b]Our specific structure also allows us to directly add up aggregate effects estimated via Equation (9) and local effects estimated via Equation (10) to obtain an estimate of total individual effects, somewhat reminiscent of the results in Wolf (2023).

## 3.1 Alternative Representations

Before turning to the details of the estimation, it is useful to give two alternative, equivalent representations of our model. Those are useful because they connect our work to frameworks that may be more familiar to the reader.

**Representation 1: A Factor Model**

We first define the vector of all idiosyncratic variables as

$$X_t = [X_t^{1\prime} \; X_t^{2\prime} \; \ldots \; X_t^{N\prime}]'$$

Then we can stack all idiosyncratic equations to arrive at the following expression:

$$X_t = X_t^{agg} \otimes \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}_{N \times 1} + \mu^X + \sum_{l=1}^{L^{reg}} \tilde{A}_l^X X_{t-l} + \sum_{l=1}^{L^{agg}} \tilde{C}_l^{agg} X_{t-l}^{agg} + B^X \eta_t + \varepsilon_t^X \tag{11}$$

where $\otimes$ denotes the Kronecker product, and $\tilde{A}_l^X$ is a sparse and block-diagonal matrix, whereas $\tilde{C}_l^{agg}$ and $B^X$ are dense matrices. Our model thus has a factor structure at the idiosyncratic level, with factors being given by current and lagged aggregate variables as well as aggregate shocks.

The second representation is a restricted VAR, which we discuss next.

**Representation 2: A Restricted VAR**

We first define the vector of all variables as

$$Z_t = [X_t^{agg\prime} X_t']'$$

Then we can stack all equations to arrive at the following expression:

$$Z_t = \mu^Z + \sum_{l=1}^{max(L^{agg}, L^{reg}, L)} A_l^Z Z_{t-l} + \underbrace{B^Z \eta_t + \varepsilon_t^Z}_{w_t^Z} \tag{12}$$

where $w_t^Z$ is the overall forecast error and $A_l^Z$ are sparse matrices. This expression is derived by plugging the aggregate dynamics from equation (9) into each idiosyncratic set of equations (10).

As mentioned above, our model imposes parsimony by not allowing for direct spillovers between regions other than through aggregate shocks and aggregate variables. It turns out that equilibrium models that have exactly this structure are prominently used in the literature. A leading example is Jones et al. (2022), where the authors develop and estimate an equilibrium model explicitly modeling dynamics at the levels of U.S. states.[4]

## 3.2 Bayesian Estimation

We estimate the model via Bayesian methods, exploiting the Gibbs sampler. We use priors so that the conditional posteriors are all known in closed form, exploiting our assumption of Gaussian shocks and making the estimation reasonably fast. Posterior approximation algorithms such as the Gibbs sampler are inherently recursive, slowing down estimation. However, as we will discuss next, the parameters for each region can be drawn in parallel, making the estimation of this model feasible even in large cross sections.

In the application section, we provide guidance on how to choose reasonable default priors that can serve as a benchmark for further exploration. This is particularly important for parameters governing the effects of shocks ($B^i$ and $B^{agg}$), as there is no standard prior choice already present in the literature. In summary, our Gibbs sampler draws from the following conditional posteriors, building on Matthes and Schwartzman (2023):

- Conditional on the parameters in the aggregate block ($\mu^{agg}$, $\{A_l^{agg}\}_{l=1}^{L}$, $B^{agg}$, $\Sigma^{agg}$) and the regional block ($\mu^i$, $\{A_l^i\}_{l=1}^{L^{agg}}$, $\{C_l^i\}_{l=1}^{L^{reg}}$, $B^i$, $\Sigma^i$ $\forall i = 1, ..., N$) $\eta_t$ can be drawn by exploiting the Kalman filter and related smoothing algorithms for linear and Gaussian systems, based on Carter and Kohn (1994). To make this step more numerically efficient, we follow Durbin and Koopman (2012) and collapse the large vector of observables into a vector with the same dimension as the structural shocks.

- Aggregate variables ($\mu^{agg}$, $\{A_l^{agg}\}_{l=1}^{L}$, $B^{agg}$, $\Sigma^{agg}$) conditional on regional variables and $\eta_t$ can be drawn using known conditional distributions (we assume Gaussian priors for $B^{agg}$).

[4] In terms of our notation, their state level dynamics can be written as:

$$A^{agg}(L)X_t^{agg} = w_t \tag{13}$$
$$A^i(L)(X_t^i - X_t^{agg}) = w_t^i \tag{14}$$

where $A^{agg}(L)$ and $A^i(L)$ are polynomials in the lag operator and $w_t$ and $w_t^i$ are one-step forecast errors at the aggregate and state level, respectively.

- Regional variables ($\mu^i$, $\{A^i_l\}_{l=1}^{L^{agg}}$, $\{C^i_l\}_{l=1}^{L^{reg}}$, $B^i$, $\Sigma^i$ $\forall i = 1, ..., N$) conditional on aggregate variables and $\eta_t$ can be drawn using known conditional distributions (we assume Gaussian priors for $B^i$). Importantly, given independent priors across $i$, we can parallelize the drawing of these parameters.

# 4 Leveraging Information from Microeconometric Identification

We now describe in detail how a researcher can use the information from microeconomic identification strategies to estimate macroeconomic effects. We adopt a Bayesian approach, which recognizes the uncertainty around the microeconomic estimates and associated identification assumptions. Our goal is to establish priors on the impact of shocks as encoded in the matrix $B^Z$, exploiting the connection between the identification of structural parameters in time series models and priors on the impact of structural shocks, as highlighted in Baumeister and Hamilton (2015). The conditions for the identification of shocks are established in the following proposition, proven in Matthes and Schwartzman (2023):

**Proposition 1 (Matthes and Schwartzman (2023))** *Consider the state-space representation implied by the VAR representation of the model (12) augmented with the trivial state equation that the unobserved shocks are the states.*[5] *The least squares projection of $\eta^k_t$ based on current and past observables (obtained using the Kalman filter) depends only on the $k^{th}$ column of $B^Z$ ($B^Z_k$), and the covariance matrix of $Z_t - E_{t-1}Z_t$, regardless of initial conditions for the state.*

The proposition gives a set of identification conditions for any given shock $\eta^k_t$. That is, identification requires knowledge of the elements of $B^Z$, which encode the effects of $\eta^k_t$ on different variables. This information can come from (imperfect)knowledge of the effects that these shocks have on idiosyncratic, micro-level variables (after controlling for time-fixed effects, as is common in the literature), or effects of the same shocks on macroeconomic variables. However, it does not require knowledge of the specific effects of other shocks.

[5] To give more detail, the observation equation would be

$$Z_t - \mu^Z - \sum_{l=1}^{max(L^{agg},L^{reg},L)} A^Z_l Z_{t-l} = w^Z_t$$

and the state equation

$$w^Z_t = [B^Z I][\eta'_t \; \varepsilon^{Z\prime}_t]'$$

The proposition allows different approaches to determine the values in $B_k^Z$. In Matthes and Schwartzman (2023) we propose using a priori measures of exposure to shocks such as output shares. Alternatively, Sarto (2024) imposes exclusion restrictions on the direct impact of particular shocks on certain variables (hence, shocks to variable 1 only affect variable 2 indirectly through its effect on variable 1 etc) that discipline the values of $B^Z$.[6]

Our current approach is to obtain identification from the assumptions and results of microeconometric studies, possibly in addition to more traditional macroeconomic identification assumptions, thus providing a bridge from those studies to macroeconomic effects. Typically, researchers study the impact of one idiosyncratic variable (which we denote $X_{1,t}^i$) on another (denoted by $X_{2,t}^i$). They estimate models of the form

$$X_{2,t}^i - X_{2,t}^{agg} = \gamma(X_{1,t}^i - X_{1,t}^{agg}) + W_t'\xi + u_t^i, \tag{15}$$

where $W_t$ is a vector of controls and $\xi$ of parameters, and $u_t$ is a residual. This is a more general version of equation (5) in the example, including controls and with the variables denoted by indices.

Equation (15) cannot usually be estimated by OLS. A major concern is that both $X_{1,t}^i$ and $X_{2,t}^i$ are caused by a similar set of local variables (including reverse causality from $X_{2,t}^i$ to $X_{1,t}^i$). A common solution for that is to construct an instrument for $X_{1,t}^i$ based on multiplying $X_{1,t}^{agg}$ by a local coefficient $\beta_i$, where $\beta_i$ is chosen as a reasonable measure of the marginal effect of changes in $X_{1,t}^{agg}$ on $X_{1,t}^i$ *net of a time-fixed effect*. Such a measure implies by construction that the instrument is relevant.
The exclusion restriction is

$$\frac{1}{TI}\sum_{i,t}\beta_i X_{1,t}^{agg} u_t^i = \frac{1}{T}\sum_t \left[X_{1,t}^{agg} \times \left(\frac{1}{I}\sum_i \beta_i u_t^i\right)\right] = cov_T\left(X_{1,t}^{agg}, cov_I(\beta_i, u_t^i)\right) = 0$$

where the last equality follows from $\frac{1}{T}\sum_t u_t^i = 0$ and $\frac{1}{I}\sum_i u_t^i = 0$.[7] $cov_I$ denotes a cross-sectional sample covariance, while $cov_T$ denotes a sample covariance computed across time periods. An implication is that, when $X_{2,t}^{agg}$ is high, shocks cannot be systematically higher or lower in units where $\beta^i$ is higher. At the core of papers in that literature is, therefore, an argument for why the cross-sectional covariance is zero or at least uncorrelated with $X_{1,t}^{agg}$.

[6] Sarto (2024) also does not use a Kalman filter for estimation, so the discussion of identification is somewhat different in that case.

[7] In particular, $1/I\sum_i \beta_i u_i = \frac{1}{I}\sum_i \beta_i \times \frac{1}{I}\sum_i u_{it} + cov_I(\beta_i, u_t^i) = cov_I(\beta_i, u_t^i)$, and $\frac{1}{T}\sum_t X_{1,t}^{agg} cov_I(\beta_i, u_t^i) = \frac{1}{T}\sum_t X_{1,t}^{agg} \times \frac{1}{T}\sum_t cov_I(\beta_i, u_t^i) + cov_T(X_{1,t}^{agg}, cov_I(\beta_i, u_t^i)) = cov_T(X_{1,t}^{agg}, cov_I(\beta_i, u_t^i))$. The last equality follows from the fact that $\frac{1}{T}\sum_t cov_I(\beta_i, u_t^i) = \frac{1}{T}\sum_{t,i}(\beta_i - \frac{1}{I}\beta_i)(u_t^i - \frac{1}{I}u_t^i) = \sum_i(\beta_i - \frac{1}{I}\beta_i)(\frac{1}{T}\sum_t u_t^i - \frac{1}{I}\frac{1}{T}\sum_t u_t^i) = 0$. Standard regularity conditions ensure that these conditions hold, at least approximately, in moderate and large samples.

For example, because the instrument only varies over time with the aggregate variable, it will be orthogonal to local shocks.[8] For aggregate shocks, the exclusion requires that, after accounting for controls, the effects of other aggregate shocks are uncorrelated with $\beta_i$, a point emphasized by Canova (2022).

Under the identification assumptions, this methodology consistently estimates $\gamma$. This estimate can be used to set priors on the effects of a shock $\eta_t^1$ that only affects $X_{2,t}^i$ through its effect on $X_{1,t}^i$. In particular, such a shock has the property that it affects $X_{2,t}^i$ by a factor $\gamma$. Moreover, estimation of $\gamma$ involved a stance that the shock affects $X_{1,t}^i$ by a factor $\beta_i$ of its effect on $X_{1,t}^{agg}$. Bringing those together, we have that the prior means for $B_{1,1}^i$ and $B_{2,1}^i$ are

$$E[B_{1,1}^i] = b_i E[B_{1,1}^{agg}]$$
$$E[B_{2,1}^i] = \gamma E[B_{1,1}^i]$$

That is, the shock to $\eta_1$ affects $X_{1,t}^i$ through its effect on the aggregate variable, and $X_{2,t}^i$ through its effect on $X_{1,t}^i$. We conclude the discussion of prior means for the idiosyncratic block with two remarks:

First, and consistent with the class of microeconomic studies that we use, the priors imply a homogeneous impact of $X_{1,t}^i$ on $X_{2,t}^i$. In the present context, this assumption is less stringent than it appears. An advantage of the Bayesian approach is that the homogeneous effect is imposed only on the prior means, but is not imposed dogmatically. The final estimation will, in general, yield heterogeneous effects that we can calculate and describe.

Second, the procedure allows for the use of estimates obtained outside the specific econometric framework and data that we use to estimate the aggregate effects. In particular, the model can be specified without all the controls needed to render the exclusion restriction valid or using estimates obtained from a local projection rather than a VAR. All that is needed is that $\gamma$ is appropriately estimated.

## 4.1 Prior on Aggregate Effect

It remains to set the effect of $\eta_1$ on $B_{1,1}^{agg}$. We start from the observation that $\eta_1$ can be normalized so that its standard deviation is equal to one. $B^{agg}$ determines the variance (and covariance) between the different aggregate variables. We, therefore, choose to set prior means on $B^{agg}$ that are consistent with those covariances. To do that, we estimate a version of our aggregate block via ordinary least squares (OLS). We then choose the prior of $B_1^{agg}$

[8] An underlying assumption is that, unlike in Gabaix and Koijen (2023), local shocks do not have an effect on aggregates.

based on the variance of the OLS residuals. The variance of the one-step ahead forecast error $\tilde{\Sigma}^{agg}$ captures the fluctuation in $B^{agg}\eta_t$ and $u_t$. To set the prior on $B^{agg}_{1,1}$, we suppose that a fraction $\theta$ of the variance of $X^{agg}_{1,t}$ is explained by the government spending shock we identify. Given that $\eta_t$ has a unit variance, our prior mean for $B^{agg}_{g,1}$ is

$$E\left[B^{agg}_{1,1}\right] = (\theta\tilde{\Sigma}^{agg}_{1,1})^{1/2} \tag{16}$$

The value of $\theta$ is chosen to maximize the marginal likelihood, but could alternatively be chosen directly as a prior hyperparameter.

As an alternative, note that $B^{agg}_{1,1}$ relates an aggregate shock to an aggregate outcome (often a policy variable, as in our example). As such, previous studies using aggregate data or calibrated equilibrium models can be used to inform the prior on $B^{agg}_{g,1}$.[9]

### 4.2 Incorporating Standard Macroeconomic Identification Schemes

As mentioned above, our model can easily incorporate more standard macroeconomic identification schemes since it has a (restricted) VAR representation. In particular, information on the sign and magnitudes of the impact effects of shocks on aggregates can be incorporated via priors on $B^Z$, similar to Baumeister and Hamilton (2015).[10] Zero restrictions can be incorporated (or at least approximated) via tight priors on specific elements of $B^Z$. This insight also provides an avenue for incorporating instruments for the macroeconomic shock itself (Mertens and Ravn, 2013; Plagborg-Møller and Wolf, 2021) by including the instrument as an aggregate variable and using zero restrictions as described in Plagborg-Møller and Wolf (2021).

## 5 Application: Revisiting Nakamura and Steinsson (2014)

Nakamura and Steinsson (2014) lever regional variation in defense spending to estimate local (or "open economy relative") government spending multipliers, which they use to inform dynamic equilibrium models. We use their data not only directly estimate to aggregate multiplier, but also infer total multipliers for each US state, which our model allows to be heterogeneous.

[9] Since we use the previous derivations to set non-degenerate priors on the impact matrices, our approach will technically only set-identify objects of interest. However, with a large cross-section of variables for which we use these priors, the amount of additional uncertainty due to having set identification is small (Amir-Ahmadi and Drautzburg, 2021; Matthes and Schwartzman, 2023).

[10] Approximate sign restrictions at longer horizons can be incorporated by specific choices on the lag coefficients in the model, as discussed in Baumeister and Hamilton (2015).

## 5.1 Data

We consider a bivariate system for both aggregate and regional blocks: $X_t^{agg} = (y_t^{agg}, g_t^{agg})'$ and $X_t^i = (y_t^i, g_t^i)'$ where $y$ and $g$ represent output and military spending, respectively. As in Nakamura and Steinsson (2014), these two variables are defined as the two-year difference of the corresponding raw variable normalized by output.

$$y_t^{agg} = \frac{Y_t^{agg} - Y_{t-2}^{agg}}{Y_{t-2}^{agg}}, \quad g_t^{agg} = \frac{G_t^{agg} - G_{t-2}^{agg}}{Y_{t-2}^{agg}}, \quad y_t^i = \frac{Y_t^i - Y_{t-2}^i}{Y_{t-2}^i}, \quad g_t^i = \frac{G_t^i - G_{t-2}^i}{Y_{t-2}^i}$$

All of the data is taken directly from Nakamura and Steinsson (2014). In particular, we use their choice of two-year differences. We thus end up with annual data spanning from 1967 to 2006 for 51 states. Capital letters denote real (deflated by national CPI), per capita variables.

## 5.2 Identification via Priors

In this section, we describe our priors, with a particular focus on those priors that are directly relevant for identifying the fiscal multiplier and that encode our identification assumptions. For standard, VAR-type parameters, we use Minnesota priors Doan et al. (1984), as is common in the literature. The priors except for the relevant entries of $B^Z$ are common across the two applications we present in this paper.

The government spending shock will be identified as the first element of $\eta_t$. Hence, the response of the aggregate variables to this shock is represented by the first column of $B^{agg}$, which we call $B_1^{agg} = (B_{y,1}^{agg}, B_{g,1}^{agg})'$ and the response of idiosyncratic variables is $B_1^i = (B_{y,1}^i, B_{g,1}^i)$, where for legibility we index row elements of each $B$ matrix by the variable letter rather than its position in the VAR.

Table 1 summarizes the prior distributions of the parameters involved in the aggregate and regional blocks, respectively.

| | Type of Distribution | Parameters |
|---|---|---|
| **Aggregate Block** | | |
| $\mu^{agg}, A^{agg}$ | Normal | Minnesota Prior |
| $B^{agg}$ (Elements related to shock of interest) | Normal | See main text |
| $B^{agg}$ (other) | Normal | Mean: 0.0, Std: 10 |
| $\Sigma^{agg}$ | Inverse Wishart | Scale: OLS<br>dof: 10 |
| **Regional Block** | | |
| $\mu^i, C^i$ | Normal | Minnesota Prior |
| $A^i$ | Normal | Mean: 0.0, Std: 0.5 |
| $B^i$ (Identified) | Normal | Regional information<br>(See main text) |
| $B^i$ (Unidentified) | Normal | Mean: 0.0, Std: 10 |
| $\Sigma^i$ | Inverse Wishart | Scale: OLS<br>dof: 10 |

Table 1: Prior Specifications for Aggregate and Regional Blocks

### 5.2.1 Priors on $B^i$

The key step in our identification methodology is to use prior information obtained from econometric studies using fixed-effects to impose priors on $B^i_{y,1}$ and $B^i_{g,1}$. To establish priors on the sensitivity of regional spending to the aggregate spending shock $B^i_{g,1}$, we adopt the two methods used by Nakamura and Steinsson to construct their instrument. First, we estimate the first-stage regression in Nakamura and Steinsson (2014),

$$g^i_t = \beta^i g^{agg}_t + \alpha_i + \gamma_t + \varepsilon^i_t, \quad i = 1, \cdots, N$$

The estimated coefficient $\beta^i$ is used to inform the prior mean of $B^i_{g,1}$ after being rescaled by the effect of the government spending shock on aggregate government spending.[11] This regression does not include the same controls as our time series model, which also controls for lags of the relevant variables. Below we discuss that our findings are robust to an alternative specification where this regression does include the same control variables. For the second specification, we use the average ratio between state spending and state output for the first

[11] We rescale by multiplying the coefficients from Nakamura and Steinsson (2014) by the prior mean $E\left[B^{agg}_{g,1}\right]$ described in Section 4 above. An alternative would be to use a hierarchical prior.

five years of the sample, a shift-share setup.[12] In both cases, the prior standard deviation for $B^i_{g,1}$ is set to half the absolute value of the prior mean. We set the prior mean of $B^i_{y,1}$ to the prior mean of $B^i_{g,1}$ multiplied by the corresponding estimate of the local multiplier in Nakamura and Steinsson (2014).[13] We again choose the prior standard deviation of $B^i_{y,1}$ to equal half the absolute value of its prior mean. As with local government spending, the intention here is that we use a prior that is informative enough to inform the local multiplier, but we do not want to make it dogmatic. The prior means for $B^i_{g,1}$ and $B^i_{y,1}$ average to zero for both specifications, consistent with our regional block specification in which we subtract $X^{agg}_t$ from $X^i_t$.[14]

### 5.2.2 Priors on $B^{agg}$

To set the prior on $B^{agg}_{g,1}$ we follow the procedure delineated in Section 4.1 and choose the prior mean for $B^{agg}_{g,1}$ so that $\eta_1$ accounts for a fraction $\theta$ of the variance of innovations to $G$, as estimated via OLS. We then choose $\theta$ to maximize the marginal likelihood of the estimated model.

The other prior parameters in $B^{agg}_1$, the prior mean of $B^{agg}_{y,1}$ and the standard deviation of both $B^{agg}_{y,1}$ and $B^{agg}_{g,1}$ are implied by our prior for the government spending multiplier. To pin down these parameters, we draw $m^{agg} = B^{agg}_{y,1}/B^{agg}_{g,1}$ one million times from the prior distribution for different values of prior hyperparameters until we hit our target moments for the fiscal multiplier – we target a median for the prior of the spending multiplier of 0.8 with a 90% interval of 0.5-1.5. This range is motivated by our reading of the existing literature – three representative examples are: Ramey (2019): "The bulk of the estimates across the leading methods of estimation and samples lie in a surprisingly narrow range of 0.6 to 1.", Nakamura and Steinsson (2018): "Estimates between 0.5 and 1.0–which is where most of the more credible estimates based on US data lie–.... ", and Barnichon et al. (2021): "Unfortunately, despite intense scrutiny the range of estimates for the government spending multiplier remains wide–between 0.5 and 2–...". The priors for the effects of other aggregate shocks on both aggregate and regional variables are Gaussian with a mean of 0 and a large standard deviation of 10. Details on priors for other parameters that are not directly relevant for identification of the structural shocks can be found in Appendix A.[15]

[12] The shift-share structure is treated as a robustness check in Nakamura and Steinsson (2014). However, as pointed out in Ramey (2020), the shift-share specification gives a larger first-stage $F$ statistic, so we find it useful to study both specifications here.

[13] The values for these multipliers are specification-specific and given in Appendix D.

[14] In the shift-share specification, we demean the coefficients obtained from Nakamura and Steinsson (2014).

[15] We generate 100,000 draws from our posterior, of which we discard the first 50,000.

## 5.3 The Aggregate Government Spending Multiplier

We use a lag length of 2 for both the aggregate and state-level blocks and include $R = 3$ aggregate shocks in our estimation.[16] To have a transparent discussion of the aggregate government spending multiplier we first give an explicit definition: In our specification, the impact multiplier equals $B^{agg}_{y,1}/B^{agg}_{g,1}$, the response of $y^{agg}_t$ to the identified spending shock divided by the response of $g^{agg}_t$ to the same shock. Before turning to our benchmark results, a useful question to answer is "How much could we learn from our aggregate data and standard time series methods alone?". If we want to use aggregate data alone, there are many macro-based identification schemes that we could use. For simplicity, and because it fits well with our benchmark specification in a way we describe below, we first estimate a VAR on our aggregate data using the same Minnesota prior that we use in our full model, order government spending first and use a simple Cholesky-type recursive identification scheme that identifies the government spending shock as the forecast error of government spending. Since we use defense spending, this is an assumption that is both reasonable and transparent.[17] The resulting 90 percent posterior interval centered at the median is $(-0.33, 5.71)$, with a median of 2.70. It is safe to say that with our annual dataset we cannot learn anything useful from aggregate data alone.[18] Table 2 instead summarizes the results for the case with the prior based on the first-stage regression in Nakamura and Steinsson (2014) and our full model. We show the prior and three posteriors, one where we do not use *any* local information (second column)[19], one where we only use information on the local response of government spending (third column), and one where we additionally use information on the local multiplier from Nakamura and Steinsson (2014) to inform the local response of output to a government spending shock (last column, our benchmark specification). With our benchmark specification, the marginal likelihood is maximized at $\theta = 1$, which means that all variation in the forecast error of government spending comes from the government spending shock, in line with the recursive identification scheme we used for the aggregate-only VAR. We can see that using identification information at the micro-level does shift multiplier (posterior median) estimates by over 10 percent when compared to our prior or the model without

[16] We provide guidance on how to pick the number of aggregate shocks in Section 6.

[17] Although, as stressed by Nakamura and Steinsson (2018), the assumption excludes the possibility of geopolitical shocks affecting both military spending and output directly

[18] We intentionally want an apples-to-apples comparison here. If a researcher only used aggregate data, it is safe to say that they would use data at a higher frequency, which we cannot do because we also want to use state-level data.
The exact identification scheme for aggregate-only VAR turns out to be less important - we find even wider posterior bands if we, for example, use a time-aggregated version of the military news series of Ramey and Zubairy (2018).

[19] When we don't use an informative prior, we set the prior standard deviation of the corresponding elements to 10.

information on local effects. This confirms that local information can play an important role in the identification of macroeconomic effects. However, using micro-level information on local effects of defense spending does not move the multiplier substantially above 1 (the probability that the aggregate multiplier is larger than 1 increases to 37 percent from our prior of 28 percent). Not incorporating information on the local effects of government spending on output leads to a somewhat smaller increase in the probability of 5 percent.

| | Prior | Posterior | Posterior | Posterior |
|---|---|---|---|---|
| | 0.80 | 0.80 | 0.85 | 0.91 |
| $m^{agg}$ | (0.38, 1.55) | (0.37, 1.53) | (0.40, 1.65) | (0.46, 1.41) |
| | [0.53, 1.18] | [0.52, 1.17] | [0.56, 1.25] | [0.63, 1.20] |
| $Prob(m^{agg} > 1)$ | 0.28 | 0.28 | 0.34 | 0.37 |
| Log MDD | | -7337.99 | -7227.25 | -7283.38 |
| $\theta$ | | 1.00 | 0.25 | 1.00 |
| Informative $B_y^i$ prior | | No | No | Yes |
| Informative $B_g^i$ prior | | No | Yes | Yes |

Table 2: Results based on Nakamura and Steinsson (2014) first-stage regression. 90% posterior bands are in parentheses, and 68% bands are in square brackets. Results with Informative prior for $B_y^i$ represent our benchmark results.

These qualitative results are robust to using the shift-share results instead to inform our prior, as shown in Table 3. Both point estimates of the aggregate multiplier and the estimated probability of the multiplier being greater than 1 are now larger, with an 18 percent increase in the probability relative to the prior or the case with uninformative local priors (as seen in Table 2)[20] This finding confirms that cross-sectional information helps to move the prior, but does not provide conclusive evidence that the aggregate multiplier is larger than 1 (or smaller than 1, for that matter).

[20]Our multiplier estimate implicitly averages over different monetary policy regimes that could have been in place during the sample, so they are not inconsistent with the takeaways in Nakamura and Steinsson (2014). Similarly, we use a linear model. If nonlinear effects are important for government spending multipliers, as argued by Barnichon et al. (2021), then again we estimate an average multiplier.

| | Posterior | Posterior |
|---|---|---|
| $m^{agg}$ | 0.90<br>(0.41, 1.71)<br>[0.59, 1.32] | 0.97<br>(0.49, 1.52)<br>[0.68, 1.29] |
| $Prob(m^{agg} > 1)$ | 0.39 | 0.46 |
| Log MDD | -7208.78 | -7292.27 |
| $\theta$ | 1.00 | 1.00 |
| Informative $B_y^i$ | No | Yes |
| Informative $B_g^i$ | Yes | Yes |

Table 3: Results based on Nakamura and Steinsson (2014) shift-share setting. 90% posterior bands are in parentheses, and 68% bands are in square brackets.

### 5.3.1 Discussion

We now highlight various alternative specifications we have explored to get a sense of how robust our results are. In particular, we are interested in robustness in two dimensions: (i) data sources/transformations, and (ii) prior specifications. Tables with the detailed results for each specification can be found in Appendix B. First, results are robust to using one year differences instead of two-year differences when computing the observables. Results are also broadly unchanged when directly computing aggregate observables as output-weighted averages of the regional variables. In terms of priors, one concern readers might have is that the Nakamura and Steinsson (2014) first-stage regression does not use the same control variables that we include in our model. To confront this possible issue, we estimate a new version of their first stage regression that does include the same variables on the right-hand side as our regional VAR block. We again find no substantial evidence of an aggregate multiplier larger than 1.

Next, we address the role that the aggregate prior plays. Loosening the aggregate prior leads to point estimates (posterior median) of the aggregate multiplier that are larger, but posterior uncertainty dramatically increases, making any meaningful statement about the aggregate multiplier impossible. For example, with a prior for aggregate multiplier centered at 0 and the 90 percent prior probability interval centered at the median going from -6 to 6, we get a posterior median of 2.19, but a 90 percent posterior band centered at the median going from -1.49 to 5.72. Why does the aggregate prior matter so much then? Is it because the regional identification information is not useful? No, the estimated government spending shock is basically the same in our benchmark and this specification (the correlation is basically 1, as shown in Appendix C), but the effect our estimated shock has on aggregate output is not tightly pinned down by data alone, so the aggregate prior plays an important role, in particular because the time series dimension is rather small in our sample. We thus find that in order to pin down aggregate effects of government spending, aggregate information is not only useful, but even crucial.

Next we graphically assess the role of two features of our benchmark prior. First, we have thus far picked $\theta$ to maximize the marginal likelihood. How much does this matter? Figure 1 gives an answer.

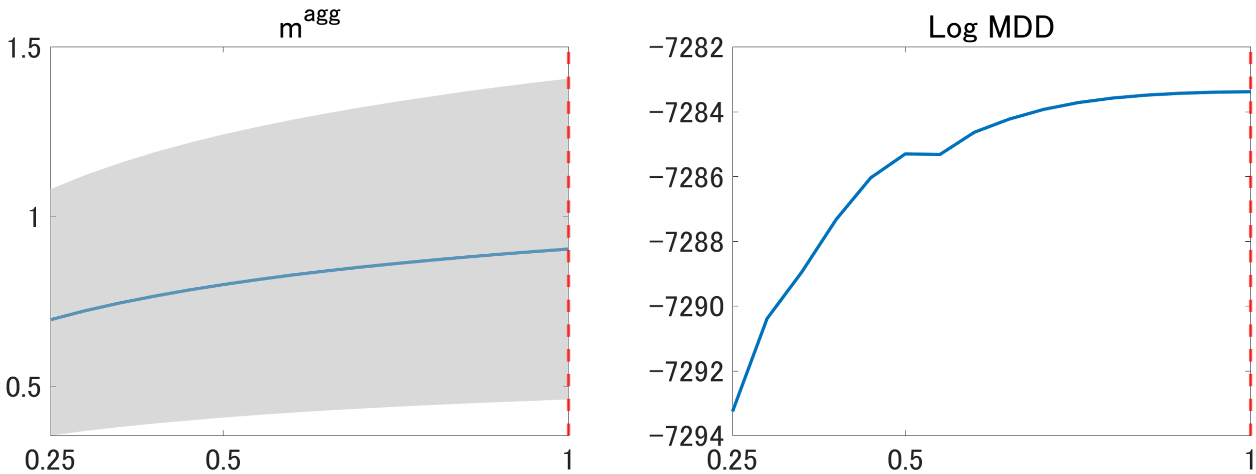


Figure 1: Changing $\theta$. Left panel plots the aggregate multiplier (median and 90 percent posterior bands), right panel plots the marginal data density estimated via method in Geweke (1999). Dashed red vertical line shows the benchmark $\theta$ value.

Although the fit of the model increases substantially with $\theta = 1$, the qualitative conclusions about the multiplier remain unchanged.

Second, we check how important local prior information is by changing the associated standard deviation. In our benchmark, we set the standard deviations for all local effects of government spending to half the absolute value of the corresponding mean. Figure 2 shows what happens when we use other values than 0.5. Again, our findings of multipliers that are smaller than unity remain.

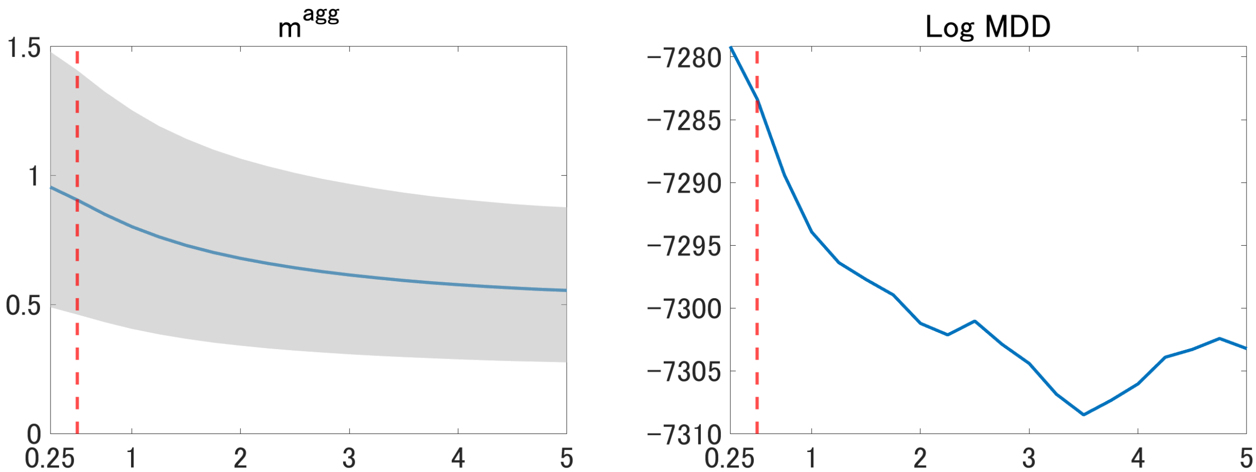


Figure 2: Changing standard deviation of prior on local effects. Left panel plots the aggregate multiplier (median and 90 percent posterior bands), right panel plots the marginal data density estimated via method in Geweke (1999). Dashed red vertical line shows the benchmark $\theta$ value.

### 5.4 The Local Government Spending Multiplier

Finally, we want to highlight the heterogeneity in the effects of government spending shocks in the United States. The open economy relative multiplier in Nakamura and Steinsson measures "the effect that an increase in government spending in one region of the union relative to another has on relative output", which can be characterized as $d(y_t^i - \bar{y}_t)/d(g_t^i - \bar{g}_t)$. In our approach, the impact open economy multiplier is thus equal to the ratio of the corresponding elements of $B^i$.[21] Moreover, we can infer the total multiplier for each state $i$ as

$$\frac{B^i_{y,1} + B_{y,1}}{B^i_{g,1} + B_{g,1}}$$

Nakamura and Steinsson (2014) estimate a homogeneous local multiplier by construction, while both our local multiplier and the total multiplier, which we will focus on here, are allowed to be heterogeneous. Figure 3 shows the estimated total multipliers as defined before for each state, along with 68 percent posterior bands. We find substantial heterogeneity, with a number of states not having any meaningful effects, whereas a substantial number of states have multipliers around or above 1.[22]

[21] Remember that we define multipliers as ratios of impulse responses.

[22] The corresponding figure with 90 percent posterior bands can be found in Appendix E. Our conclusions are robust to using these 90 percent posterior bands instead.

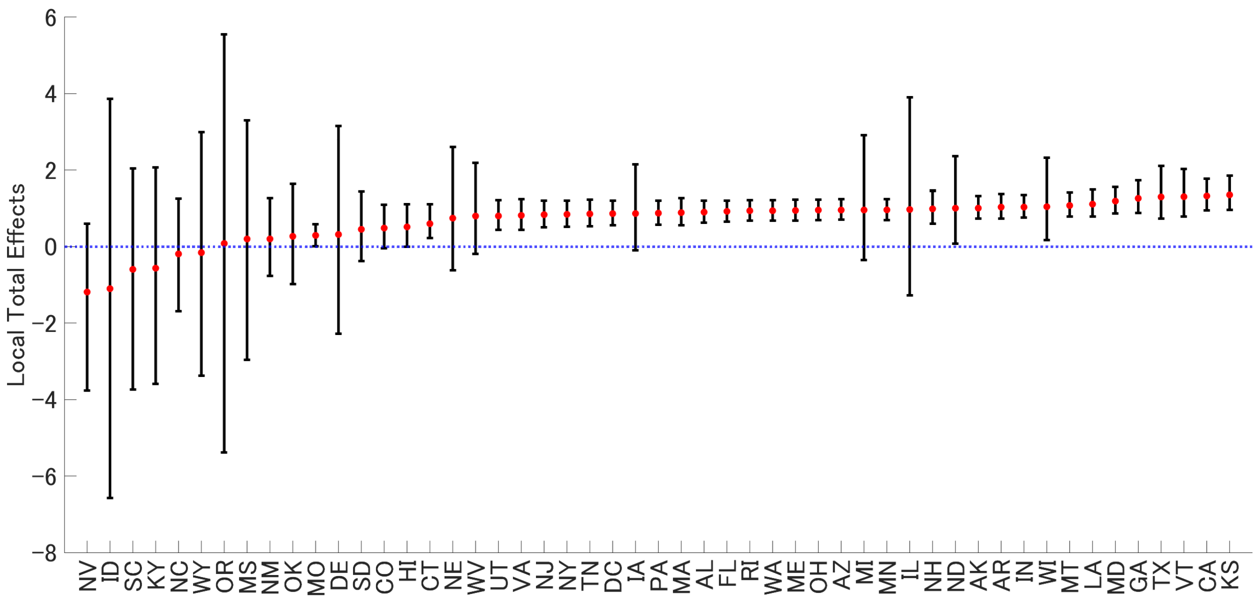


Figure 3: Median of Regional Total Effects with 68% Posterior Interval

# 6 Monte Carlo

To assess the performance of our algorithm across different sample sizes $N$ and $T$, we conduct Monte Carlo simulation exercises using the posterior median from the baseline estimation (reported in Table 2) as the data generating process (DGP)[23]. The prior of $B^{agg}_{g,1}$ and $B^{agg}_{y,1}$ are largely uninformative - they are centered at the truth for convenience but the prior standard deviations are set to be large (a value of 10). The prior mean of $B^{i}_{g,1}$ is equal to the truth, and its standard deviation is half of the absolute value of mean. The corresponding mean for the local impact on output is set to our benchmark estimate of the local multiplier (1.43) times the mean of $B^{i}_{g,1}$, as in our empirical application.[24] The assumption that the regional prior is centered on the truth reflects our view that our identification assumptions are valid, but there is substantial uncertainty. Since our aggregate prior here is uninformative, all identification comes from the regional information. We choose the prior distributions of the rest of the parameters to be the same as in the empirical application.

[23]When the number of regions $N$ is different from the one in the empirical exercise (51), we randomly generate the states using the following procedure: Let $n = \lfloor N/51 \rfloor$. For the 1st to $51n$-th states, we repeat the 51 states in the empirical benchmark for $n$ times. For the $(51n + 1)$-th to $N$-th states, we randomly draw the states from the empirical benchmark without duplication. For example, when $N = 138$, two sets of the US states are included in the 1st to 102nd states, and the remaining 34 states are drawn randomly from the observed 51 states. The selection of the states is fixed across simulations with the same choice of $(T, N)$.

[24]The standard deviation of the local output effect is set to the absolute value of the mean.

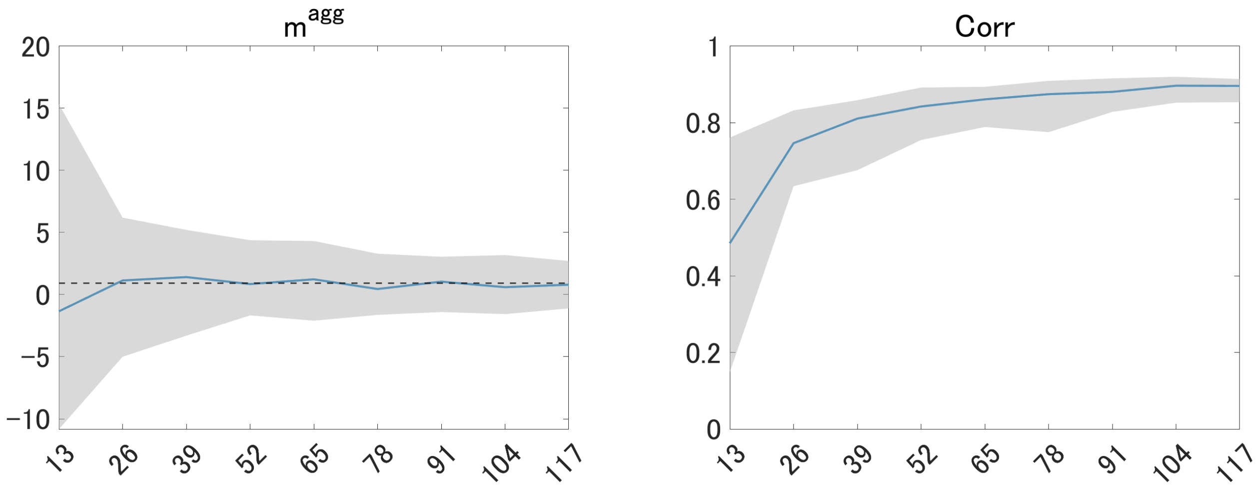


Figure 4: Sensitivity to the number of time series observations $T$ ($N = 51$)

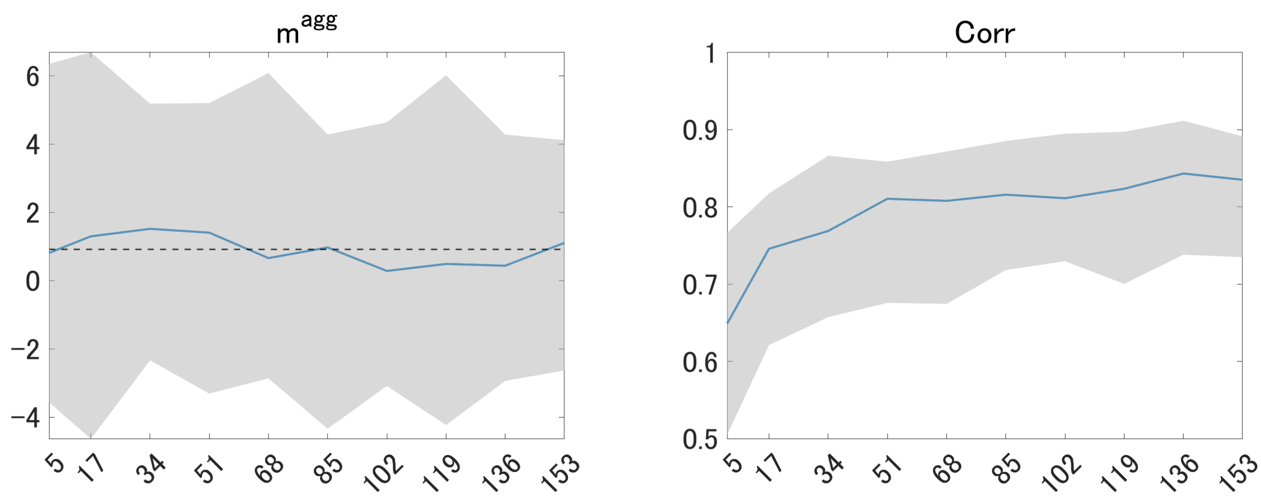


Figure 5: Sensitivity to the number of cross-sectional units $N$ ($T = 39$).

Figures 4 and 5 explore the sensitivity to the sample length and the size of the cross-section respectively. The solid blue line represents the median of the posterior medians from 48 simulations along with the 90% interval constructed from those 48 medians. The dashed line represents the true value of the parameters, which is equal to the prior mean. The top right panel reports the correlation between the true and identified (posterior median) aggregate shocks. Overall, we can see that adding longer time series helps, whereas increasing the cross-section has no effect, meaning that 51 states already provide all the cross-sectional variation that can be exploited in this application. This also mirrors our discussion in Section 5, where the limited time-series dimension of our sample limits what one can learn about the aggregate multiplier in the absence of an informative aggregate prior. Relative to our benchmark findings in the Nakamura and Steinsson (2014) application, the uninformative nature of the aggregate prior in this Monte Carlo results in substantial uncertainty/dispersion of estimates across Monte Carlo samples, as can be seen in the uncertainty bands constructed from the medians across our 48 samples.

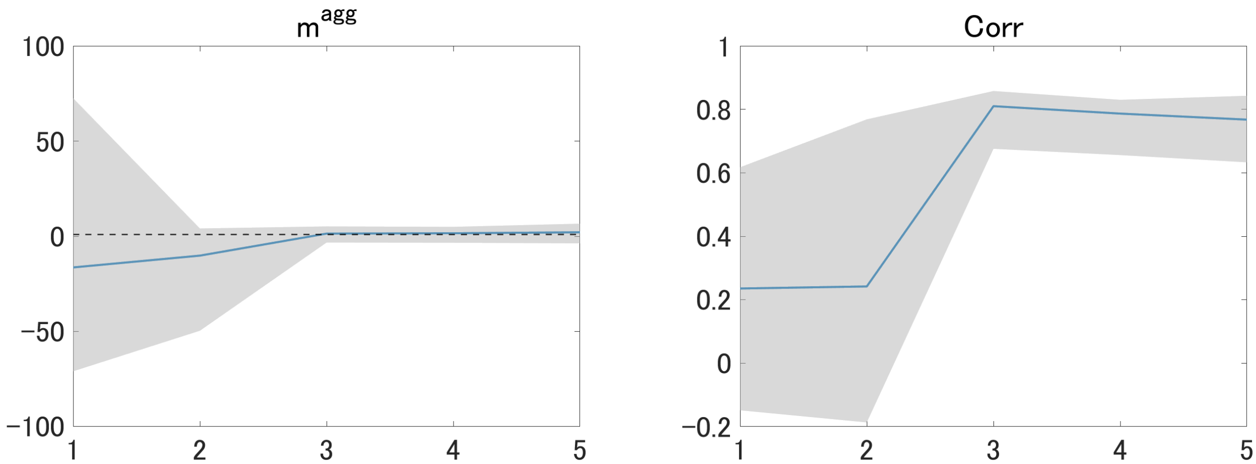


Figure 6: Sensitivity to $R$ ($R^{true} = 3$)

To investigate the sensitivity to the assumption on the number of aggregate shocks $R$, we generate the sample with $R^{true} = 3$ and estimate the model with different assumptions on $R$. Figure 6 plots the outcome of this exercise. We can see that once the correct number of shocks is included, increasing the number of shocks further has no effect. This result can be used as a guide for empirical applications: Researchers should choose to increase $R$ until the results do not change anymore when $R$ is increased further.

To see how well our procedure recovers the aggregate shock of interest, we pick one particular simulation and compare the posterior distribution of the identified aggregate shock with the truth. With the same sample size as the empirical application (Figure 7), the extracted shock series keeps track of the truth very well. The true shock series is mostly within the posterior bands even though the bands are tight.

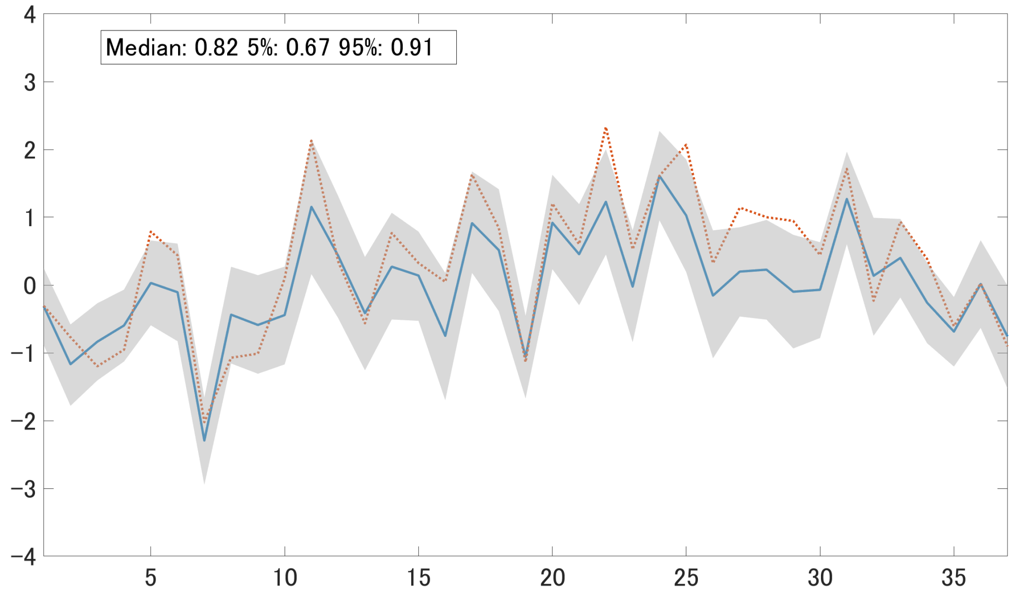


Figure 7: One simulated shock series along with the estimated shock and 90 percent posterior bands for that sample. $(T, N) = (39, 51)$. Legend gives percentiles of the distribution of correlation between the true shock and estimated shock.

# 7 Conclusion

We have presented an econometric framework that can jointly leverage identification strategies from the applied micro toolkit, and identification assumptions from the macro/time series literature and as such exploits both time series and cross-sectional variation to identify aggregate macroeconomic effects as well as total idiosyncratic effects of identified shocks. This stands in contrast with results obtained using standard applied micro tools that estimate time-fixed effects and as such take out the macroeconomic effects we are interested in.

Using a well known application on government spending multipliers, we highlight how aggregate identification information is still needed to obtain sharp estimates of many objects of interest, whereas others (such as the shock itself) can be obtained without informative priors on aggregate effects of this shock.

# References


Amir-Ahmadi, P. and T. Drautzburg (2021, January). Identification and inference with ranking restrictions. *Quantitative Economics 12*(1), 1–39.

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

Barnichon, R., D. Debortoli, and C. Matthes (2021, 06). Understanding the Size of the Government Spending Multiplier: It's in the Sign. *The Review of Economic Studies 89*(1), 87–117.

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

Baumeister, C. and J. D. Hamilton (2015, September). Sign Restrictions, Structural Vector Autoregressions, and Useful Prior Information. *Econometrica 83*(5), 1963–1999.

Baumeister, C. and J. D. Hamilton (2023). A Full-Information Approach to Granular Instrumental Variables. Technical report.

Borusyak, K., P. Hull, and X. Jaravel (2021, 06). Quasi-Experimental Shift-Share Research Designs. *The Review of Economic Studies 89*(1), 181–213.

Canova, F. (2022). Should we trust cross sectional multiplier estimates? *Journal of Applied Econometrics, forthcoming*.

Canova, F. and G. D. Nicolo (2002, September). Monetary disturbances matter for business fluctuations in the G-7. *Journal of Monetary Economics 49*(6), 1131–1159.

Carter, C. K. and R. Kohn (1994, 09). On Gibbs sampling for state space models. *Biometrika 81*(3), 541–553.

Chodorow-Reich, G. (2019, May). Geographic cross-sectional fiscal spending multipliers: What have we learned? *American Economic Journal: Economic Policy 11*(2), 1–34.

Chodorow-Reich, G. (2020). Regional data in macroeconomics: Some advice for practitioners. *Journal of Economic Dynamics and Control 115*.

Christiano, L. J., M. Eichenbaum, and C. L. Evans (1999). Monetary policy shocks: What have we learned and to what end? Volume 1 of *Handbook of Macroeconomics*, pp. 65–148. Elsevier.

Doan, T., R. Litterman, and C. Sims (1984). Forecasting and conditional projection using realistic prior distributions. *Econometric Reviews 3*(1), 1–100.

Durbin, J. and S. J. Koopman (2012). *Time Series Analysis by State Space Methods*. Oxford University Press.

Faust, J. (1998). The robustness of identified VAR conclusions about money. *Carnegie-Rochester Conference Series in Public Policy 49*, 207–244.

Gabaix, X. and R. S. Koijen (2023). Granular instrumental variables. *Journal of Political Economy forthcoming*.

Geweke, J. (1999). Using simulation methods for bayesian econometric models: inference, development,and communication. *Econometric Reviews 18*(1), 1–73.

Goldsmith-Pinkham, P., I. Sorkin, and H. Swift (2020, August). Bartik Instruments: What, When, Why, and How. *American Economic Review 110*(8), 2586–2624.

Jones, C., V. Midrigan, and T. Philippon (2022). Household leverage and the recession. *Econometrica 90*(5), 2471–2505.

Matthes, C. and F. Schwartzman (2023). The consumption origins of business cycles: Lessons from sectoral dynamics. Working paper, Richmond Fed.

Mertens, K. and M. O. Ravn (2013, June). The Dynamic Effects of Personal and Corporate Income Tax Changes in the United States. *American Economic Review 103*(4), 1212–1247.

Moll, B. (2021). From cross-section to aggregates: the "missing intercept problem". Technical report.

Nakamura, E. and J. Steinsson (2014, March). Fiscal stimulus in a monetary union: Evidence from us regions. *American Economic Review 104*(3), 753–92.

Nakamura, E. and J. Steinsson (2018, August). Identification in macroeconomics. *Journal of Economic Perspectives 32*(3), 59–86.

Plagborg-Møller, M. and C. K. Wolf (2021, March). Local Projections and VARs Estimate the Same Impulse Responses. *Econometrica 89*(2), 955–980.

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

Ramey, V. A. (2019, May). Ten years after the financial crisis: What have we learned from the renaissance in fiscal research? *Journal of Economic Perspectives 33*(2), 89–114.

Ramey, V. A. (2020, May). *Comment on "What Do We Learn From Cross-Regional Empirical Estimates in Macroeconomics?*, pp. 232–241. University of Chicago Press.

Ramey, V. A. and S. Zubairy (2018). Government spending multipliers in good times and in bad: Evidence from us historical data. *Journal of Political Economy 126*(2), 850–901.

Sarto, A. P. (2024). Recovering Macro Elasticities from Regional Data. Technical report.

Uhlig, H. (2005, March). What are the effects of monetary policy on output? Results from an agnostic identification procedure. *Journal of Monetary Economics 52*(2), 381–419.

Wolf, C. K. (2023, August). The Missing Intercept: A Demand Equivalence Approach. *American Economic Review 113*(8), 2232–2269.

# Appendix For "Estimating The Missing Intercept"

## A More Information on Priors

The parameters other than $B^{agg}$ and $B^i$ are set following standard practice in the VAR literature. The scale of the inverse Wishart distributions for the covariance matrix of residuals is chosen on the basis of the OLS estimation of a VAR with the same variables. To be more precise, we estimate (9) and (10) without acknowledging the factor structure in the forecast errors and set the estimated $\tilde{\Sigma}^{agg}$ and $\tilde{\Sigma}^i$ $(i = 1, \cdots, N)$ as a prior mean for the covariance matrix of the residuals. We use a small number of degrees of freedom (10) so that this prior is not very informative.

Our prior for the aggregate response of government spending to a government spending shock is parameterized via $\theta$ (which we choose to maximize the marginal likelihood in the government spending application using the Geweke (1999) approach) as follows:

$$E\left[B^{agg}_{g,1}\right] = (\theta\tilde{\Sigma}^{agg}_{2,2})^{1/2} \tag{A-1}$$

where we assume that the aggregate government spending variable is ordered second in the VAR estimated via OLS.

### A.1 More on Minnesota Prior

**Prior Mean.** The prior mean is 0 for all coefficients other than the ones associated with own first lags, which are 1.

**Prior Variance.** The prior variance in the Minnesota prior is a diagonal matrix, where the variance of the coefficient in the $i$-th equation associated with the $l$-th order lag of $j$-th variable is given by

$$\begin{cases} \left(\frac{\phi_0}{h(l)}\right)^2 & i = j \\ \left(\phi_0 \frac{\phi_1}{h(l)} \frac{\sigma_j}{\sigma_i}\right)^2 & i \neq j \\ (\phi_0\phi_2)^2 & \text{for constants and exogenous variables} \end{cases}$$

where $\sigma_i$ and $\sigma_j$ are the square roots of the $(i, i)$ and $(j, j)$ elements in the error variance matrix. We obtain the estimate of the error variance matrix by applying OLS to (9) and

(10) without factors. The prior hyperparameters are set as $\phi_0 = 0.2$, $\phi_1 = 0.5$, $\phi_2 = 10^5$, and $h(l) = l$.

# B More Results for Government Spending Application

## B.1 1-Year Differenced Data

| | Prior | Posterior (First-Stage) | Posterior (Shift-Share) |
|---|---|---|---|
| | 0.80 | 0.74 | 0.86 |
| $m^{agg}$ | (0.38, 1.55) | (0.38, 1.15) | (0.43, 1.34) |
| | [0.53, 1.18] | [0.52, 0.98] | [0.60, 1.14] |
| $Prob(m^{agg} > 1)$ | 0.28 | 0.14 | 0.30 |
| Log MDD | | -6991.57 | -7020.17 |
| $\theta$ | | 0.50 | 0.65 |
| Prior Type | | First-Stage | Shift-Share |

Table A-1: Observables based on one-year differences. 90% posterior bands are in parentheses, and 68% bands are in square brackets.

## B.2 Output-Weighted Aggregate Data

| | Prior | Posterior (First-Stage) | Posterior (Shift-Share) |
|---|---|---|---|
| | 0.80 | 0.90 | 0.96 |
| $m^{agg}$ | (0.38, 1.55) | (0.46, 1.39) | (0.49, 1.49) |
| | [0.53, 1.18] | [0.63, 1.18] | [0.67, 1.27] |
| $Prob(m^{agg} > 1)$ | 0.28 | 0.36 | 0.44 |
| Log MDD | | -7273.47 | -7289.41 |
| $\theta$ | | 1.00 | 1.00 |
| Prior Type | | First-Stage | Shift-Share |
| Informative $B_y^i$ | | Yes | Yes |

Table A-2: Aggregate observables are output-weighted averages of regional data. 90% posterior bands are in parentheses, and 68% bands are in square brackets.

## B.3 Alternative First-Stage Regression

| | Prior | Posterior | Posterior |
|---|---|---|---|
| | 0.80 | 0.87 | 0.96 |
| $m^{agg}$ | (0.38, 1.55) | (0.41, 1.68) | (0.49, 1.49) |
| | [0.53, 1.18] | [0.57, 1.28] | [0.67, 1.27] |
| $Prob(m^{agg} > 1)$ | 0.28 | 0.36 | 0.44 |
| Log MDD | | -7227.72 | -7288.26 |
| $\theta$ | | 0.30 | 1.00 |
| Informative $B_y^i$ | | No | Yes |

Table A-3: First-stage regression now includes same controls as our baseline model. 90% posterior bands are in parentheses, and 68% bands are in square brackets.

## B.4 Looser Aggregate Prior

| | Prior | Posterior | Posterior |
|---|---|---|---|
| | 0.77 | 0.95 | 0.97 |
| $m^{agg}$ | (-0.11, 3.34) | (-0.09, 3.95) | (0.11, 1.90) |
| | [0.26, 1.68] | [0.34, 2.04] | [0.45, 1.52] |
| $Prob(m^{agg} > 1)$ | 0.37 | 0.47 | 0.48 |
| Log MDD | | -7226.98 | -7284.03 |
| $\theta$ | | 0.25 | 1.00 |
| Informative $B_y^i$ | | No | Yes |

Table A-4: 90% posterior bands are in parentheses, and 68% bands are in square brackets.

### B.5 Even Looser Aggregate Prior

| | Prior | Posterior | Posterior |
|---|---|---|---|
| $m^{agg}$ | -0.00 | 28.64 | 2.19 |
| | (-6.31, 6.32) | (-138.93, 199.42) | (-1.49, 5.73) |
| | [-1.82, 1.82] | [4.94, 74.11] | [0.03, 4.31] |
| $Prob(m^{agg} > 1)$ | 0.25 | 0.85 | 0.71 |
| Log MDD | | -7229.34 | -7292.88 |
| $\theta$ | | 0.30 | 1.00 |
| Informative $B_y^i$ | | No | Yes |

Table A-5: 90% posterior bands are in parentheses, and 68% bands are in square brackets.

## C Correlation Between Estimated Government Spending Shocks

We compute the correlation of the posterior medians of identified spending shocks from three specifications shown in right columns of Tables 2, A-4, and A-5.

| | Baseline | Looser | Even Looser |
|---|---|---|---|
| Baseline | 1 | 1.0000 | 0.9998 |
| Looser | 1.0000 | 1 | 0.9998 |
| Even Looser | 0.9998 | 0.9998 | 1 |

Table A-6: Correlation between estimated posterior median shock series.

## D Local Multiplier Estimates Used to Set Priors

| | First Stage | Shift share |
|---|---|---|
| Two-year | 1.43 | 2.48 |
| One-year | 0.69 | — |
| Alternative First-Stage Regression | 0.63 | — |

## E Total Regional Multiplier With 90 Percent Bands

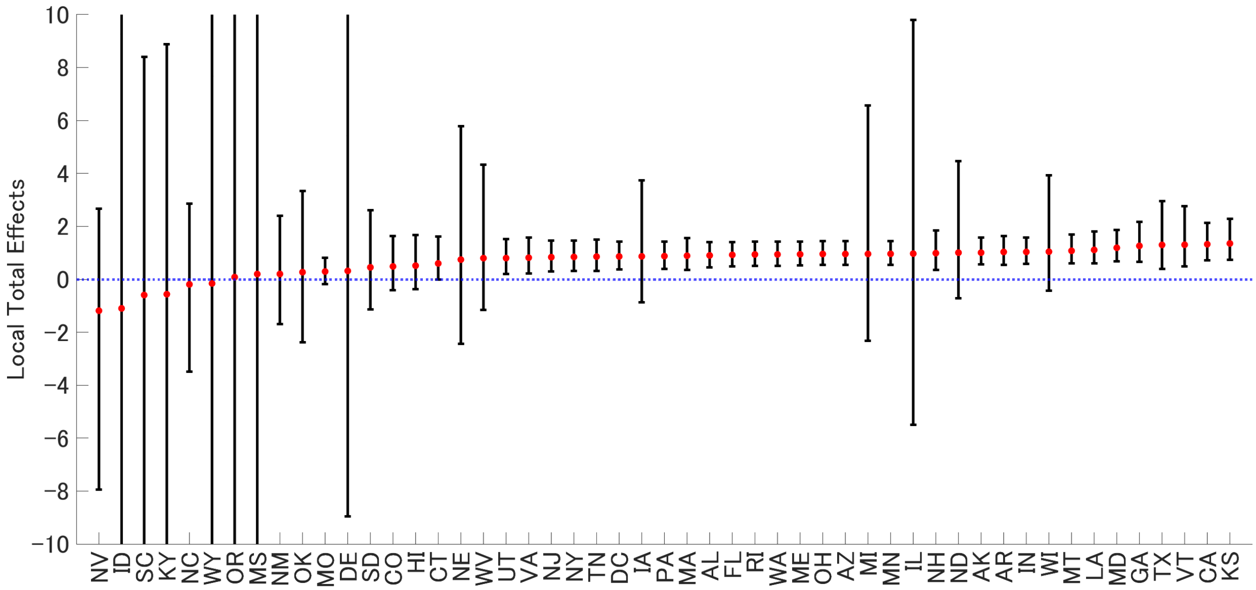


Figure A-1: Median of Local Total Effects with 90% Posterior Interval