Estimating The Missing Intercept

Christian Matthes, Naoya Nagasaka, and Felipe Schwartzman*

August 24, 2025

Abstract

Cross-sectional data have proven increasingly useful to answer questions of interest to macroeconomists. However, their use also often leads to the 'missing intercept' problem in which aggregate general equilibrium effects and policy responses are averaged out. We present a statistical approach that leverages identification in cross-sectional data to jointly identify aggregate and idiosyncratic effects in a panel framework, unifying identification approaches in cross-section and time-series. We then apply our methodology to study government spending multipliers (Nakamura and Steinsson, 2014) and wealth effects from stock returns (Chodorow-Reich et al., 2021).

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

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

*University of Notre Dame, Indiana University & Federal Reserve Bank of Richmond. Contact information: cmatthes@nd.edu, naonagas@iu.edu, felipe.schwartzman@rich.frb.org. We would like to thank Mark Watson as well as seminar participants in Berlin, EPGE-FGV, 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 increasingly relies on panel datasets encompassing heterogeneous regions, households, or firms. The cross-sectional variation in these datasets allows researchers to apply microeconometric tools to credibly identify the effects of policies or shocks. Those studies often derive their identification from differencing out confounding aggregate variation common across observational units.

However, as is well recognized, the flip side of eliminating common sources of variation through differencing or time-fixed effects is that these methods can only uncover the idiosyncratic or local effects of policy changes *net* of those aggregate effects.[1] Those tools are therefore inadequate to *directly* answer questions about aggregate effects of policies. A typical response in the literature is to use estimates based on cross-sectional variation to calibrate fully specified dynamic equilibrium models.[2] While informative, those strategies provide estimates of aggregate effects that depend on the specifics of the structural model.

We propose a methodology that combines microeconometric methods that provide sharp identification at the idiosyncratic level with flexible time-series methods that estimate aggregate effects without imposing strong cross-equation restrictions. In particular, we show how to incorporate estimates of local effects obtained from cross-sectional data into a time series model that jointly describes the evolution of aggregate and local economic outcomes. In effect, we unify the time-series and cross-sectional approaches to identification, allowing for the simultaneous use of variation in both dimensions to sharpen estimates. Furthermore, our approach makes it possible to combine the microeconometric approach with identification assumptions more commonly 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; Baumeister and Hamilton, 2015) on the impact of shocks, as well as instruments for aggregate shocks (Mertens and Ravn,

[1]See Moll (2021) for an influential exposition of the problem

[2]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.

2013; Plagborg-Møller and Wolf, 2021).

As an example, consider the following stylized description of the estimation of fiscal multipliers based on cross-state variation, as in Nakamura and Steinsson (2014). Suppose it is known that national government spending shocks increase local spending by more in Virginia than in Wisconsin by a known amount. By comparing the response of the two states to aggregate government spending fluctuations, one can therefore estimate the regional impact of government spending shocks. This method differences out confounding aggregate effects while yielding estimates of the local government spending multiplier. It does not, however, provide per se an estimate of the economy-wide, or aggregate, multiplier.

We develop a method to estimate the aggregate multiplier using the same identification assumptions employed in those studies. The only additional assumptions required are standard in macroeconomic models and the time series literature: (i) that fundamental structural shocks are uncorrelated and (ii) that the comovement across units in the panel data are well captured by a factor model, with one of the factors representing an aggregate government spending shock. Identification is then possible because the assumptions and estimates used for the cross-section constrain the local effects of the aggregate policy shock. Within the example, a government spending shock increases local government spending in Virginia by a certain amount more than in Wisconsin. Also, it increases output in these two regions proportionately by the amount implied by the estimated multiplier. Thus, observation of such comovement in the data provides information about the shock path. As in VAR-IV models, knowledge of the effects of a shock provides enough information to identify its trajectory. This identified aggregate shock then provides exogenous variation required to infer the aggregate output multiplier. As a bonus, this same aggregate shock also refines estimates of local effects, allowing for a richer degree of heterogeneity in effects than may have been originally possible.

From a technical standpoint, the method leverages the dimensionality reduction afforded by factor models. Those well-established models are designed to best summarize the comovement in large data sets using a relatively small number of variables. By imposing this flexible yet parsimonious structure, the factor framework enables us to identify a large number of objects from a relatively small number of constraints.

In practice, we perform the estimation within a single Bayesian model. In particular, we follow Baumeister and Hamilton (2015) and impose identification restrictions as “soft” priors, rather than dogmatic constraints. 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 apply the method to two prominent studies: Nakamura and Steinsson's (2014) analysis of fiscal multipliers across U.S. states and Chodorow-Reich et al.'s (2021) estimates of the effect of stock market wealth on local economies. One key difference between the two studies is in the size of the dataset used. Nakamura and Steinsson (2014) use annual data at the state level, whereas Chodorow-Reich et al. (2021) uses quarterly frequency data at county-level. For the first application we find that, while the methodology provides a sharp estimate of the fiscal shock it provides relatively little information on the aggregate multiplier. In contrast, when using the richer data in Chodorow-Reich et al. (2021), we find large and significant effects of stock market wealth on aggregate employment and the wage bill.

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. 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. Baumeister and Hamilton (2023) extend this idea to VAR settings. 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 describes through a simplified example how microeconometric studies deliver estimates of local effects and how they can be levered to estimate aggregate effects. Section 3 provides a general statement of the assumptions and methodology used. It also provides general proposition describing how full identification of local and aggregate effects can be obtained from a small number of linear restrictions. We use this general characterization to discuss linkages to previous work. Section 4 lays out in detail the general time series model that we use to lever local effects for identification of aggregate effects. Sections 5 and 6 provide two applications of our approach, building on Nakamura and Steinsson (2014) and Chodorow-Reich et al. (2021), respectively. Section 7 concludes. The online appendix includes Monte Carlo evidence on the performance of our approach, details of the estimation procedure and robustness exercises for the applications.

2 Identification and the Missing Intercept: an Example

To set the stage, we consider the impact of fiscal shock in a New Keynesian model with I regions in a monetary union indexed $i \in \{1, ..., I\}$. Using this example, we describe what objects can identified from the cross-sectional variation. The example is purposefully simplified for ease of exposition, and the insights generalize readily for other environments.

Each region $i \in \{1, \cdots, I\}$ is inhabited by a representative household with population mass $N_i = 1/I$ and featuring separable preferences over consumption and leisure.

$$U(\{c_{i,t}\}_{t=0}^{\infty}, \{\ell_{i,t}\}_{t=0}^{\infty}) = E \sum_{t=0}^{\infty} e^{\xi_{i,t}} \beta^t \left[\frac{(c_{i,t})^{1-\sigma} - 1}{1-\sigma} - \chi \frac{\ell_{i,t}^{1+\eta}}{1+\eta} \right], \tag{1}$$

where $\ell_{i,t}$ is employment per worker in region i at time t and $c_{i,t}$ is their consumption and $\xi_{i,t}$ is a region-specific discount-rate shock.

Households maximize (1) subject to the intertemporal budget constraint

$$P_{i,t} c_{i,t} + \frac{B_{i,t}}{1+r_t} = P_{i,t} w_{i,t} \ell_{i,t} + B_{i,t-1} - T_{i,t},$$

where, for each region i, time t, $P_{i,t}$ is the nominal price level, $w_{i,t}$ is the real wage, $B_{i,t}$ are one period nominal bonds, r_t is the (net) nominal rate of interest on bonds purchased at time t and $T_{i,t}$ are net taxes or transfers to residents of region i. Regions are in a monetary union, so that the nominal interest rate r_t is the same for all i.

Households maximization implies the labor supply relation and Euler equation:

$$\chi \ell_{i,t}^{\eta} c_{i,t}^{\sigma} = w_{i,t}, \quad e^{\xi_{i,t}} c_{i,t}^{-\sigma} = \beta(1+r_t) E_t \frac{e^{\xi_{i,t+1}}}{1+\pi_{i,t+1}} c_{i,t+1}^{-\sigma}$$

where $\pi_{t+1} = P_{t+1}/P_t - 1$ is inflation rate between periods t and $t+1$.

Production follows the conventional structure in New Keynesian models. In every location, firms produce one of a continuum of varieties (indexed v). They hire household labor at the real wage rate $w_{i,t}$, and transform it one for one into output $y_{i,t}(v) = \ell_{i,t}(v)$, with $\ell_{i,t}(v)$ the quantity of labor used to produce variety v in region i at time t and local labor market clearing implying $\int_v \ell_{i,t}(v)dv = \ell_{i,t}$. Firms have monopoly over the particular variety of goods that they produce. They sell those varieties to household that aggregate them into their consumption basket according to the Dixit-Stiglitz preferences, $c_{i,t} = \left[\int y_{i,t}(v)^{\frac{\epsilon-1}{\epsilon}}\right]^{\frac{\epsilon}{\epsilon-1}}$.

Firms set nominal prices for extended periods as in Calvo (19..). Since labor is their only input, they set prices to be a discounted average of future expected nominal wage rates.

In this simple example, regions are not connected through trade, so local production is used either as part of local consumption or local government expenditure ($g_{i,t}$).[3] The local resource constraints are

$$y_{i,t} = c_{i,t} + g_{i,t}$$

Taking a linear approximation around the steady-state and doing the usual derivations we obtain for each region i, time period t:

$$\pi_{i,t} = \kappa \left(\eta \frac{\tilde{y}_{i,t}}{y_i} + \sigma \frac{\tilde{c}_{i,t}}{c_i} \right) + \beta E_t \pi_{i,t+1}, \tag{2}$$

$$\frac{\tilde{c}_{i,t}}{c_i} - \tilde{\xi}_{i,t} = -\frac{1}{\sigma} \left(r_t - E_t \pi_{i,t+1} \right) + E_t \left(\frac{\tilde{c}_{i,t+1}}{c_i} - \tilde{\xi}_{i,t+1} \right) \tag{3}$$

$$\tilde{y}_{i,t} = \tilde{c}_{i,t} \ + \ \tilde{g}_{i,t}, \tag{4}$$

where $\tilde{y}_{i,t} = y_{i,t} - y_i$ denote the deviation of output $y_{i,t}$ from its steady-state value, y_i and analogously for other variables. The first and second equations are simply the regional versions of the Phillips Curve and consumption Euler equations. The last equation is a linear approximation for the local resource constraints.

The regions belong to a monetary union, which means that a single monetary

[3] Allowing for trade-linkages would allow for different wedges between local and aggregate effects - we leave this out for simplicity.

authority sets a common nominal interest rate. The monetary authority targets average inflation across regions

$$r_t = \phi \sum_i \frac{1}{I} \pi_{i,t}$$

where $\phi > 1$, so that the Central Bank follows the Taylor principle: if inflation rises, the nominal interest rate rises more than one-for-one.

In the system of equations above, the exogenous variables are the deviations of local government spending $\tilde{g}_{i,t}$ and the discount rate shock $\tilde{\xi}_{i,t}$ from their respective steady-state values. We solve for a rational expectations equilibrium where $\tilde{g}_{i,t}$ and $\tilde{\xi}_{i,t}$ are iid over time, so that $E_t \tilde{g}_{i,t+1} = E_t \tilde{\xi}_{i,t+1} = 0$ at all t irrespective of the state of the economy. Given that the system lacks any lagged dependent variable, deviations of endogenous variables from steady-state are also iid with mean zero in equilibrium.

Assuming that the steady-state consumption share of output is the same in all locations, we have that $c_i/y_i = C/Y \equiv \bar{c}$, where Y and C are, respectively, the steady-state values of aggregate consumption and output. Then, using the iid assumption we can solve for the percent deviation in local output:

$$\frac{\tilde{y}_{i,t}}{y_i} = \frac{\tilde{g}_{i,t}}{y_i} + \bar{c} \left(\tilde{\xi}_{i,t} - \frac{1}{\sigma} r_t \right), \tag{5}$$

That is, output increases with local government spending $g_{i,t}$ and local discount shock $\xi_{i,t}$, but declines with the national interest rate r_t set by monetary policy. With some additional steps, one can find the interest rate as a function of aggregate government spending and discount factors, defined as $\tilde{G}_t \equiv \sum_i \tilde{g}_{i,t}$ and $\tilde{\Xi}_t \equiv \sum_i \frac{1}{I} \tilde{\xi}_{i,t}$:

$$r_t = \frac{\eta \phi \kappa \sigma}{\sigma + (\sigma + \eta \bar{c}) \phi \kappa} \frac{\tilde{G}_t}{Y} + \frac{(\sigma + \eta \bar{c}) \phi \kappa \sigma}{\sigma + (\sigma + \eta \bar{c}) \phi \kappa} \tilde{\Xi}_t.$$

It follows that the interest rate rises with aggregate government spending and with the discount rate shock. Plugging this back into equation (5) yields an expression for local output as a function of local and aggregate forcing terms:

$$\frac{\tilde{y}_{i,t}}{y_i} = \underbrace{m^{\text{local}}}_{=1} \frac{\tilde{g}_{i,t}}{y_i} - \theta_G \frac{\tilde{G}_t}{Y} + \bar{c}\left(\tilde{\xi}_{i,t} - \theta_\Xi \tilde{\Xi}_t\right), \tag{6}$$

where m^{local} is the *local* multiplier, and $\theta_G \equiv \frac{\phi\,\kappa\,\eta C}{(1+\phi\,\kappa)\,\sigma\,Y\,+\,\phi\,\kappa\,\eta C}$, and $\theta_\Xi = \frac{\phi\kappa(\sigma Y+\eta C)}{\sigma Y+\phi\kappa(\sigma Y+\eta C)}$ where both θ_G and θ_Ξ lie between 0 and 1.

Comparing equations (5) and (6) one can see that, while local output depends on local government spending directly, it also depends on aggregate government spending indirectly through its impact on the interest rate. The multiplier of local government spending, m^{local}, is equal to 1. Fixing local government spending, local output *declines* with aggregate government spending. This is because as aggregate government spending increases, interest rates also increase, reducing local consumption and output.

To obtain the response of aggregate output to aggregate government spending, multiply both sides of (6) by y_i, add them up across i, and divide by Y:

$$\frac{\tilde{Y}_t}{Y} = \underbrace{m^{agg}}_{=(1-\theta_G)\,\in\,[0,1]} \frac{\tilde{G}_t}{Y} + \bar{c}(1-\theta_\Xi)\tilde{\Xi}_t, \tag{7}$$

where m^{agg} is the aggregate government spending multiplier. It follows that this spending multiplier is positive but smaller than 1. The aggregate multiplier incorporates the positive direct effects of local government spending on local output and the negative general equilibrium effect through interest rates.

To close the model, we now describe the determination of $\tilde{g}_{i,t}$ and $\tilde{\xi}_{i,t}$ as functions of exogenous shocks. In particular, local discount rate have aggregate and idiosyncratic components

$$\tilde{\xi}_{i,t} = \gamma_i \eta_t^\Xi + \epsilon_{i,t}^\xi$$

where η_t^Ξ is an iid aggregate shock that affects all regions and $\epsilon_{i,t}^\xi$ are local iid shocks with mean zero. For simplicity, we make assumptions reminiscent of a "law of large numbers", so that $\sum \frac{1}{I}\epsilon_{i,t}^\xi = 0$ and $\sum \frac{1}{I}\gamma_i = 1$. It then follows that the aggregate discount rate $\tilde{\Xi}_t$ varies with the exogenous aggregate shock η_t^Ξ, $\tilde{\Xi}_t = \eta_t^\Xi$

Local and aggregate government spending have similar structure, but also respond to discount shocks:

$$\frac{\tilde{G}_t}{Y} = -\alpha^{agg}\eta_t^{\Xi} + \eta_t^G, \quad \frac{\tilde{g}_{i,t}}{y_i} = \beta_i \frac{\tilde{G}_t}{Y} - \alpha\epsilon_{i,t}^{\xi} + \epsilon_{i,t}^G \tag{8}$$

Equation (8) indicates that aggregate government spending leans against aggregate demand shocks η_t^{Ξ} at the rate α^{agg}, and is otherwise subject to exogenous fluctuations from shocks η_t^G. Those exogenous shocks can incorporate, for example, geopolitical considerations driving military spending. The second equation indicates a similar leaning of local government spending against local demand shocks, $\epsilon_{i,t}^{\xi}$ while also allowing for exogenous fluctuations ϵ_{it}^G. Furthermore, local government expenditures $g_{i,t}$ are a function of aggregate expenditures $\tilde{G}_t$. Thus, for example, a geopolitical shock η_t^G increasing aggregate government expenditure $\tilde{G}_t$ has a local effect $\beta_i\eta_t^G$.

2.1 Estimating Multipliers

With the framework in hand, we can describe different methods for estimating the fiscal multipliers m^{local} and m^{agg}, and the conditions under which they are valid.

Time series approach The more traditional approach uses time-series data of the aggregate variables to estimate m^{agg} in equation (7). From inspection, it is clear why equation (7) cannot normally be estimated simply by OLS regression of $\tilde{Y}_t/Y$ on $\tilde{G}_t/Y$: the residual term $\bar{c}(1-\theta)\tilde{\Xi}_t = \bar{c}(1-\theta)\eta_t^{\xi}$ correlates with the independent variable through the feedback rule in (8). In other words, government spending reacts to offset aggregate demand shocks that move consumption, so that a regression of output on government spending understates the multiplier.

As discussed in detail in Nakamura and Steinsson (2018a), the time-series literature takes two main approaches to identification. One is to impose restrictions on the contemporaneous impact of shocks on aggregate variables to separate out η_t^G from η_t^{ξ}. For example, when using a Cholesky decomposition in a VAR, the researcher may assume that η_t^ξ does not affect $\tilde{G}_t$ contemporaneously ($\alpha^{agg} = 0$).

Assumption TS-OLS: $E\left[\frac{\tilde{G}_t}{Y}(\frac{\tilde{Y}_t}{Y} - m^{agg}\frac{\tilde{G}_t}{G})\right] = 0$,

where E is the expected value over all time-periods t. The second approach is to search for instruments that correlate with η_t^G but not η_t^ξ. For example, when estimating fiscal multipliers, military spending provides a popular source of exogenous variation. That is, given an instrument Z_t correlated with $\tilde{G}_t/Y$, the time-series approach requires

Assumption TS-IV: $E\left[Z_t(\frac{\tilde{Y}_t}{Y} - m^{agg}\frac{\tilde{G}_t}{G})\right] = 0$.

As with any instrumental variable approach, the main difficulty is that such instruments may be contentious. For example, as Nakamura and Steinsson (2018a) point out, military spending may not be valid instruments if geopolitical shocks simultaneously affect military spending and output through channels unrelated to government spending.

Cross-sectional approach A second approach is to leverage cross-sectional variation. In particular, in the context of the fiscal multiplier, Nakamura and Steinsson (2014) use state-level variation to obtain the local effect of government spending shocks. The focus on cross-sectional variation allows one to control for common sources of variation through the use of time fixed-effects. For our example, subtracting the aggregate $\tilde{Y}_t$ equation (7) from the local $\tilde{y}_{i,t}$ equation (6) the estimating equation becomes

$$\Delta\frac{\tilde{y}_{i,t}}{y_i} = \Delta\frac{\tilde{g}_{i,t}}{y_i} + \bar{c}\Delta\tilde{\xi}_{i,t}, \tag{9}$$

where we use Δ to denote the difference between local and aggregate variable, so that $\Delta\frac{\tilde{y}_{i,t}}{y_i} = \frac{\tilde{y}_{i,t}}{y_i} - \frac{\tilde{Y}_t}{Y}$ and so on.

Taking out time-effects is helpful if the aggregate discount rate shock η_t^ξ affects demand in all regions similarly ($\gamma_i = 1\ \forall i$). The differencing of aggregate variables eliminates the confounding variation in η_t^ξ from equation (6). To see this most clearly, write the regressor $\Delta\tilde{g}_{i,t}/y_i$ and the error term $\Delta\tilde{\xi}_{i,t}$ in equation (9) as a function of shocks:

$$
\begin{aligned}
\Delta \frac{\tilde{g}_{i,t}}{y_i} &= (\beta_i - 1)\eta_t^G - (\beta_i - 1)\alpha\eta_t^\xi - \alpha_i \epsilon_{i,t}^\xi + \epsilon_{i,t}^G \\
\Delta \tilde{\xi}_{i,t} &= (\gamma_i - 1)\eta_t^\xi + \epsilon_{i,t}^\xi
\end{aligned}
$$

With $\gamma_i = 1$, η_t^ξ drops from the residual $\Delta\tilde{\xi}_{i,t}$, eliminating a source of correlation with the regressor $\Delta\frac{\tilde{g}_{i,t}}{y_i}$. OLS is still biased, though, if local fiscal authorities react to local discount rate shocks ($\alpha^i \neq 0$), in which case the local discount-rate shock $\xi_{i,t}^\xi$ will lead to co-movement in $\Delta\frac{\tilde{g}_{i,t}}{y_i}$ and $\Delta\tilde{\xi}_{i,t}$.

To handle endogeneity stemming from local shocks, the cross-sectional approach often relies on a "shift-share" type instrument, constructed by multiplying a measure of $\beta_i - 1$ by $\tilde{G}_t/Y$. Given equation (8) the instrument $Z_{i,t} = (\beta_i - 1)\frac{\tilde{G}_t}{Y}$ can be expressed in terms of shocks as

$$
Z_{i,t} = (\beta_i - 1)\left(-\alpha\eta_t^\xi + \eta_t^G\right)
$$

Note that, by construction, the instrument does not depend on local shocks $\epsilon_{i,t}^\xi$, allowing it to capture variation that is orthogonal to that. With this instrument, the IV estimator of the multiplier is

$$
\widehat{m}^{cross-section} = \frac{cov\left(\Delta\frac{\tilde{y}_{i,t}}{y_i}, (\beta_i - 1)\frac{\tilde{G}_t}{Y}\right)}{cov\left(\Delta\frac{\tilde{g}_{i,t}}{y_i}, (\beta_i - 1)\frac{\tilde{G}_t}{Y}\right)} = m^{\text{local}} - \frac{\bar{c}\alpha\sigma_\Xi^2}{\alpha^2\sigma_\Xi^2 + \sigma_G^2}\frac{cov(\gamma_i, \beta_i)}{var(\beta_i)} \qquad (10)
$$

where σ_G^2 and σ_Ξ^2 are, respectively, the variances of η_t^G and η^Ξ.

Note that the estimator requires that β^i has to be heterogeneous across regions. Otherwise the instrument lacks cross-sectional variance and cannot be relevant.

Assumption CS-1: *β^i is either known or has been estimated, and $var(\beta^i) > 0$*

The estimator recovers the *local* multiplier $m^{\text{local}} = 1$ in the special case mentioned above where local discount factors are affected in the same way by the aggregate shock η_t^ξ ($\gamma_i = 1\ \forall i$). More generally, it is valid if those local effects are uncorrelated with the local exposure to aggregate government spending, $cov(\gamma_i, \beta_i) = 0$. The condition can alternatively be written as

Assumption CS-2: $E[G_t E_t[\beta_i(y_{i,t} - m^{local} g_{i,t})]] = 0$

where the expectation operation in the interior conditions on time t information and takes the average across cross-sectional units i, while the outer expectation takes the average across time-periods. To address the possibility that this may not hold, studies following this methodology may add controls to absorb variation in local exposures.

While useful for many purposes, the procedure cannot recover the aggregate multiplier $m^{agg} = m^{local} - \theta_G$. This is because in the methodology above, the term $\theta_G G_t$ is absorbed by time effects and cannot be estimated. This constitutes the *missing intercept problem.*

Combining cross-sectional and time-series variation The methodology laid out in this paper seeks to combine cross-sectional and time-series variation to identify the aggregate multiplier $1 - \theta_G$. We write the system in vector form as

$$\begin{aligned} w_t^i &= B^i \eta_t + \tilde{\epsilon}_t^i \text{ for } i \in \{1, ..., I\} \\ w_t^{agg} &= B^{agg} \eta_t \end{aligned} \tag{11}$$

where $w_t^i = \{\Delta \tilde{g}_{i,t}/y_i, \Delta \tilde{y}_{i,t}/y_i\}^T$, $w_t^{agg} = \{\tilde{G}_t/Y, \tilde{Y}_t/Y\}^T$ collect the observed variables, $\eta_t = \{\eta_t^G, \eta_t^\Xi\}^T$ and $\{\epsilon_{i,t}^g, \epsilon_{i,t}^\xi\}$ collect the shocks. The matrices B^i and B^{agg} collect the parameters and, in particular,

$$B_{\eta^G,Y}^{agg} = m^{agg} B_{\eta^G,G}^{agg}, \quad B_{\eta^G,Y}^i = m^{local} B_{\eta^G,G}^i, \tag{12}$$

where $B_{\eta^G,Y}^{agg}$ is the entry of B^{agg} capturing the impact of η^G on Y etc. In words, the multipliers refer to the relative impact of the government spending shock η_t^G on output as compared to government spending itself and analogously for the other elements.[4]

[4]More generally, we can write the aggregate block as

From (11), it is apparent that Δg_{it}, Δy_{it}, $\tilde{G}_t/Y$ and $\tilde{Y}_t/Y$ have a *factor structure.* That is, they are determined by a few aggregate shocks η_t^G and η_t^Y that simultaneously affect values in several regions, with region-specific loadings ($B^i_{\eta^G,G}$, $B^i_{\eta^G,Y}$ etc), as well as by idiosyncratic shocks $\tilde{\epsilon}^y_{it}$ and $\tilde{\epsilon}^g_{it}$.

Identification is possible given information on $B^i_{G,G}$ and $B^i_{Y,G}$. The logic is similar to the one found in the IV-based identification of shocks in structural VARs. First, econometric theory makes clear that given a large enough panel one can can estimate the systematic part of $\bar{w}_t$, given by

$$\bar{w}^i_t = B^i \eta_t$$

As the number of regions I increases, $\bar{w}^i_t$ is consistently estimated. For the heuristic discussion that follows we will assume that we have an accurate measure of $\bar{w}^i_t$.

The key identification challenge is that the observables $\bar{w}^i_t$ are functions of unobserved shocks η_t. It is the same challenge faced in the structural VAR literature. As in that literature, the challenge can be met by adding structure to B^i.[5]

We can borrow that structure from the assumptions made in the cross-sectional

---

$$\frac{\tilde{G}_t}{Y} = B^{agg}_{\eta^G,G}\eta^G_t + B^{agg}_{\eta^\Xi,G}\eta^\xi_t, \quad \frac{\tilde{Y}_t}{Y} = B^{agg}_{\eta^G,Y}\eta^G_t + B^{agg}_{\eta^\Xi,Y}\eta^\xi_t$$

where, given the example, $B^{agg}_{\eta^G,G} = 1$, $B^{agg}_{\eta^\Xi,G} = -\alpha$, $B^{agg}_{\eta^G,Y} = 1 - \theta^G$ and $B^{agg}_{\eta^\Xi,Y} = -(1-\theta^G)\alpha + \bar{c}(1-\theta_\Xi)$. And the regional block by

$$\Delta\frac{\tilde{g}_{i,t}}{y_i} = B^i_{\eta^G,G}\eta^G_t + B^i_{\eta^\Xi,G}\eta^\Xi_t + \tilde{\epsilon}^i_{Y,t}, \quad \Delta\frac{\tilde{y}_{i,t}}{y_i} = B^i_{\eta^G,Y}\eta^G_t + B^i_{\eta^\Xi,Y}\eta^\Xi_t + \tilde{\epsilon}^i_{G,t}$$

where $B^i_{\eta^G,G} = (\beta_i - 1)$, $B^i_{\eta^\Xi,G} = -\alpha(\beta_i-1)$, $B^i_{\eta^G,Y} = m^{local}(\beta_i-1)$, $B^i_{\eta^\Xi,Y} = -(\beta_i-1)\alpha + \bar{c}(\gamma_i - 1)$ and the $\tilde{\epsilon}^G_{i,t}$ and $\tilde{\epsilon}^Y_{i,t}$ are reduced form residuals satisfying $\tilde{\epsilon}^i_{G,t} = -\alpha_i\epsilon^\xi_{i,t} + \epsilon^G_{i,t}$ and $\tilde{\epsilon}^i_{Y,t} = (\bar{c}-\alpha)\epsilon^\xi_{i,t} + \epsilon^G_{i,t}$.

[5] To see the identification problem, note that given any conformable orthogonal matrix Q,

$$\bar{w}^i_t = \underbrace{B^i Q^{-1}}_{\tilde{B}^i}\underbrace{Q\eta_t}_{\tilde{\eta}_t} \ \forall i,$$

so that if η_t contains more than one element we can construct infinitely many shock series $\tilde{\eta}_t$ and shock loadings $\tilde{B}^i$ that will deliver the same $\bar{w}^i_t$ in sample and have the same first and second moments in population. The same arguments hold for the aggregate series w^{agg}_t.

approach:

Assumption P-1: *Assumption CS-1 holds so that the shock* η_g^G *has heterogeneous impact on* $g_{i,t}$ *proportional to* β_i*:* $B^i_{\eta^G,G}/B^{agg}_{\eta^G,G} = \beta_i$

Assumption P-2: *Assumption CS-2 holds, so that asymptotically* $B^i_{\eta^g,Y}/B^i_{\eta^g,G} \equiv m^{local} = \widehat{m}^{cross-section}$

Assumptions P-1 and P-2 are the same needed for the construction, relevance and validity of the instrument used in the cross-sectional approach. Thus, the conditions that allow for identification of the local multiplier also allow for identification of η^G. Those conditions are, in fact, more than enough for identification of η_t^G. This is because the collection of all B^i's has more elements than the number of shocks collected in the vector η_t.[6].

One may be skeptical of assumptions P-1 and P-2 as stated. For that reason, we implement those assumptions not as hard restrictions on parameters, but as "soft", Bayesian priors. What this means in practice is that we assign probabilities to different values of β_i and $\widehat{m}^{cross-section}$ to express uncertainty about their true values, and report the distribution of estimates associated with those different values.

The last step is to recognize that a credible estimate of η_t^G can be used as an instrument Z_t conforming to assumption TS-IV, as it is orthogonal to the residual in the aggregate Y_t equation. This is, of course, again analogous to structural VARs - once one has enough restrictions in an impact matrix to identify aggregate shocks, one can likewise estimate the effects of those shocks on endogenous variables.

3 Identification in a Generalized Framework

We now consider a general abstract model economy, which we can apply to a wide range of economic environments. The main assumption we retain from above is that the model allows for a linear representation in which all endogenous variables can be expressed as functions of exogenous shocks, and that the model admits a factor structure. In particular, we consider the generalized version of (11)

[6]Technically, the covariance of $\bar{w}_t^i$ has rank equal to the dimensionality of η_t

$$\Delta w_t^i - E_{t-1}\Delta w_t^i = B^i \eta_t + \epsilon_t^i, \text{ for } i \in \{1, ..., I\} \quad (13)$$
$$w_t^{agg} - E_{t-1} w_t^{agg} = B^{agg}\eta_t + \epsilon_t$$

where now Δw_t^i is a $K \times 1$ vector of unit-specific variables expressed in deviations from aggregates and w_t^{agg} is a $N \geq K$ vector of aggregate variables, including aggregations of the variables in w_t^i. As before, η_t is a $R < IK$ dimensional vector of iid shocks affecting aggregates and all units, whereas ϵ_t^i are unit specific shocks, potentially correlated across variables within but not across units.[7]

Thus, for each of I idiosyncratic units, the model tracks 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. Furthermore, variables are now written in deviation from their previously expected values, acknowledging that parts of the vector w_t^i may be forecastable and concentrating attention on the innovations.

A model as in (13) can typically be derived as the reduced form of a linearized structural model. To determine the objects of interest, we split the vector w_t^i (and the corresponding parts of w_t^{agg}) into subvectors collecting outcome variables (y_t^i, y_t^{agg}) and policy variables (g_t^i, g_t^{agg}), so that $w_t^i = \{y_t^i, g_t^i\}^T$ and $w_t^{agg} = \{y_t^{agg}, g_t^{agg}\}^T$. Those labels stem from the structural model economy underlying (13), and incorporate the assumption that policy makers choose g_t^i and affect y_t^i as a result.[8] The main focus is on the effect of policy variables (g_t^i, g_t^{agg}) on the outcome variables (y_t^i, y_t^{agg}).

Such a causal impact can in principle be assessed within the model given a reduced

[7]Relative to the example in the previous section, we now also introduce the corresponding N -dimensional residual vector ϵ_t shock at the aggregate level. This vector allows us to have a unified approach to inference, independently of whether $R = K$, $R < K$, or $R > K$. Otherwise, the case $R < K$ would imply stochastic singularity at the aggregate level, and the case $R = K$ would mean that, conditional on the impact matrix B^{agg}, regional data would not play a role in determining the estimated time series of structural shocks.

[8]In particular, the 13 can be derived from a linearized structural model that specifies values for innovations to outcome variables $\{y_t^i - E_{t-1}y_t^i\}$ variables as functions of policy treatments $\{g_t^i - E_{t-1}g_t^i\}$ and aggregates. Using hats to denote innovations (so that $\widehat{y}_t^i = y_t^i - E_{t-1}y_t^i$, etc),

form assignment function $\Delta g^i_{k,t} = B^i_{\eta^G_k, g_k}\eta^G_{k,t}$, where $\eta^G_{k,t}$ is an aggregate, iid policy shock affecting policy variables k at time t, and $B^i_{\eta^G_k, g_k}$ is the element of B^i determining the effect of $\eta^G_{k,t}$ on $\Delta g^i_{k,t}$. The main identification challenge is that the shocks in η^G_t are generally unobserved and need to be estimated.

Denote the element of B^i encoding the effect of a shock to the k^{th} element of g^i and g^{agg} to the j^{th} element of y^i by $B^i_{\eta^G_k, y_j}$ and analogously for effects on elements of g^i, y^{agg} or g^{agg}. With this notation and assumptions in place we can define the main object of interest as

1. Local multiplier of the j^{th} element of y^i to the k^{th} element of g^i in unit i: $m^{i,local}_{g_k,y_j} = \frac{B^i_{\eta^G_k, y_j}}{B^i_{\eta^G_k, g_k}}$,

2. Aggregate multiplier of the j^{th} element of y^{agg} to the k^{th} element of g^{agg} :

$$H^i_y \widehat{y}^i_t = A^i_{g,y}\widehat{g}^i_t + A^i_{G,y}\widehat{g}^{agg}_t + A^i_{Y,y}\widehat{y}^{agg}_t + \tilde{B}^i_y \eta_t + \epsilon^i_{y,t} \text{ for } i \in \{1, ..., I\}$$

where H^i_y is a matrix with the same dimensionality as $\widehat{y}^i_t$ and $A^i_{g,y}, A_{G,y}, A^i_{Y,y} and B^i_y$ are conformable to the rest. The model also features relationships establishing the assignment of treatments to units, potentially also as functions of outcomes and aggregates:

$$\widehat{g}^i_t = A^i_{y,g}\widehat{y}^i_t + A^i_{G,g}\widehat{g}^{agg}_t + A^i_{Y,g}\widehat{y}^{agg}_t + \tilde{B}^i_g \eta_t + \tilde{\epsilon}^i_{g,t},$$

As discussed in the text, the effects are determined in terms of the effects of elements of η_t that enter the second block of equations but not the first.

The model also features aggregation relations and a "law of large numbers",

$$\frac{1}{I}\sum_i \widehat{y}^i_t = y^{agg}_t, \quad \frac{1}{I}\sum_i \widehat{g}^i_t = g^{agg}_t, \quad \frac{1}{I}\sum_i \tilde{\epsilon}^{g,i}_t = \frac{1}{I}\sum_i \tilde{\epsilon}^{y,i}_t = 0$$

Those equations constitute a linear system of equations that, given appropriate invertibility conditions, can be solved for $\widehat{y}^i_t$, $\widehat{g}^i_t$, $\widehat{g}^{agg}_t$ and $\widehat{y}^{agg}$ as a function of shocks. By applying the "law of large numbers" one can derive an expression as in 13, up to the ε_t term in the aggregate equations which we include to facilitate the econometric implementation (see footnote 7 above for a discussion).

Note, in particular, that although the expression in 13 is expressed in terms of deviations from aggregates, it can be derived without assuming that $A^i_{Y,y} = A^i_{G,g} = 1$ are equal to 1, since innovations to the aggregate variables are spanned by the aggregate shocks η_t. We thank Andres Sarto for bringing up this point.

$$m^{agg}_{g_k,y_j} = \frac{B^{agg}_{\eta^G_k,y_j}}{B^{agg}_{\eta^G_k,g_k}}$$

3. Shock to the k^{th} element of g^{agg}: $\eta_{k,t}$

The objective is to obtain estimates for $m^{agg}_{g_k,y_j}$ and $m^{i,local}_{g_k,y_j}$ given enough restrictions on the column of B^i's associated with η^G_k. The proposition below provides sufficient conditions for point identification of the full vector of reduced effects of $\eta_{k,t}$ on both policy and outcome variables, whether local or aggregate:

Proposition 1 (Point identification of a single $IK+N$-dimensional loading column from $R < IK+N$ linear restrictions)**.** *Consider the static R-factor model $Z_t = B\eta_t + \epsilon_t$ with $Z_t \in \mathbb{R}^{IK+N}$, $B \in \mathbb{R}^{(IK+N)\times R}$ and $\eta_t \overset{i.i.d.}{\sim} \mathcal{N}(0, I_R)$, independent of ϵ_t. In the context of equation system (13), $Z_t = [w^{agg\prime}_t w^{1\prime}_t \dots w^{I\prime}_t]'$. Fix a known matrix $Q \in \mathbb{R}^{m\times(IK+N)}$ (with $m \geq R$) and suppose we know*

$$Q\,B^{(k)} \;=\; c \in \mathbb{R}^m,$$

where $B^{(k)}$ is the k^{th} column of B. Define $M := QB \in \mathbb{R}^{m\times K}$.

If $\operatorname{rank}(M) = R$ *at the true B (i.e., M has full column rank), then $B^{(k)}$ is point identified: for any orthogonal H such that $(\tilde{B}, \tilde{\eta}_t) = (BH', H\eta_t)$ also satisfies $\tilde{B}^{(k)} = c$, one must have $\tilde{B}^{(k)} = B^{(k)}$.*

Proof. Let $\tilde{B} = BH'$, so $\tilde{B}^{(k)} = BH'e_k$, where e_k is a conformable selection vector. Using the condition $Q\tilde{B}^{(k)} = QB^{(k)} = c$, we obtain

$$QBH'e_k \;=\; QBe_k \quad\Longleftrightarrow\quad M(H'e_k - e_k) \;=\; 0.$$

Since $\operatorname{rank}(M) = R$ (full column rank), $\ker(M) = \{0\}$, hence $H'e_k = e_k$. Therefore $\tilde{B}^{(k)} = B^{(k)}H'e_k = Be_k = B^{(k)}$, proving point identification. □

By taking the factor model in the proposition to include all aggregate and unit-level equations, and $B^{(k)}$ to include the reduced form effects of a particular shock of interest, the proposition states that one can obtain values for the whole $IK+N$-dimensional

$B^{(k)}$ vector by imposing $R << IK + N$ restrictions. One can then recover both the full vector of local effects and the aggregate effect, since effects can be constructed from elements in $B^{(k)}$. This is a benefit of the factor structure, as it reduces the comovement in the system to a relatively small number of factors.

The restrictions can take several forms. At its simplest, they can be assignments of values to elements of $B^{(k)}$. More subtly, they can apply to ratios of pairs of elements. Thus, assumptions such as P-1 or P-2 in the example can be imposed on the unit-level elements of $B^{(k)}$ to obtain the elements associated with the aggregate equations. Other possibilities include restrictions on aggregate effects themselves, or on averages of unit-level effects (so long as enough of those averages are available). 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.

As a consequence of allowing for identification of the full $KI + N$ vector of effects from $R < KI + N$ restrictions, the proposition implies that one may have more than enough restrictions at hand to allow for identification. In practice, however, those restrictions may by only be known imprecisely. One way to systematically incorporate the information encoded in a large number of restrictions while retaining some skepticism about each one of them is to impose those restrictions "softly", as priors in a Bayesian framework as in Baumeister and Hamilton (2018).

The conditions for the identification of the shock η_t^G are established in the following proposition, proven in Matthes and Schwartzman (2023):

Proposition 2 (Matthes and Schwartzman (2023)). *Consider the state-space representation implied by the system in proposition 1 augmented with the trivial state equation that the unobserved shocks are the states.*[9] *The least squares projection of* η_t^k

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

$$Z_t = w_t^Z$$

and the state equation

$$w_t^Z = [B\ I][\eta_t'\ \epsilon_t']'$$

based on current and past observables (obtained using the Kalman filter) depends only on the k^{th} column of B ($B^{(k)}$), and the covariance matrix of $Z_t - E_{t-1}Z_t$, regardless of initial conditions for the state.

We now turn to additional objects of interest and connect those to what is estimated in the literature:

Impulse Response Functions The multipliers above only apply to innovations, but in a more general dynamic setup one may be interested into multipliers calculated at different time intervals, that is, one may be interested in

$$m^{i,\text{local}}_{g_k,y_j,h} = \frac{\frac{\partial(y^i_{j,t+h} - E_{t-1}y^i_{j,t+h})}{\partial\eta^G_{k,t}}}{\frac{\partial(g^i_{k,t+h} - E_{t-1}g^i_{k,t+h})}{\partial\eta^G_{k,t}}}, \quad \text{and} \quad m^{\text{agg}}_{g_k,y_j,h} = \frac{\frac{\partial(y^{agg}_{j,t+h} - E_{t-1}y^{agg}_{j,t+h})}{\partial\eta^G_{k,t}}}{\frac{\partial(g^{agg}_{k,t+h} - E_{t-1}g^{agg}_{k,t+h})}{\partial\eta^G_{k,t}}}$$

Much of the literature relies on local projection methods, which impose their own set of constraints. In particular, as pointed out by Canova (2022), the literature often assumes that unit-level effects are homogeneous. In the next section, we use instead a VAR-based approach where we allow for unit-level heterogeneity in the responses.

Partial and General Equilibrium Effects Wolf (2023) examines a class of structural consumption-savings models, where individual consumption of individual i satisfies a non-linear relationship of the general form

$$c^i_t = c(Y_t, \eta^g_t, \tilde{\epsilon}_{y,t})$$

where Y_t is a vector of economy-wide of endogenous aggregates, η_t is either government spending or a shock, and $c()$ is the function mapping exogenous shocks and aggregate variables into c^i_t.

Wolf (2023) decomposes the effects of given shock η^g into a partial and a general equilibrium effect

$$\frac{dc_t^i}{d\eta_t^g} = \underbrace{\frac{\partial c}{\partial \eta_t^g}}_{\text{partial equilibrium}} + \underbrace{\frac{\partial c}{\partial Y}\frac{\partial Y}{\partial \eta_t^g}}_{\text{general equilibrium}}$$

The first term is the partial equilibrium effect, capturing the direct effect of the shock if all other feedbacks are kept constant, and the second effect is a general equilibrium effect, capturing the effect of the shock through those equilibrium feedbacks.

Wolf (2023) shows that one can estimate the partial equilibrium effect using the cross-sectional approach if we modify exposures to the shock η_t^g, so that, translating to the terminology in the current framework, it follows that

$$m^{local} = \underbrace{\frac{\partial c}{\partial \eta_t^g}}_{\text{partial equilibrium}}, \quad \text{and } m^{agg} \simeq \underbrace{m^{local}}_{\text{partial equilibrium}} + \underbrace{m^{agg} - m^{local}}_{\text{general equilibrium}}$$

The missing intercept problem is then one of finding the general equilibrium component. Wolf (2023) establishes conditions for the estimation of the missing intercept using a two-part approach. In particular, he shows that if private and public spending shocks have symmetric effects on equilibrium variables, then one can use aggregate public spending shocks to estimate general equilibrium effects.

Local spillovers Chodorow-Reich (2020) distinguishes between local, aggregate, and "all regions" effects. The local and aggregate effects correspond, respectively, to m^{local} and m^{agg} in our example. The "all regions" effect calculates the effect of a local shock to single region i aggregated across all regions, $i \in \{1, ..., I\}$. This is generally not the same as the local multiplier because of potential spillovers between regions through trade or other channels. As shown by Chodorow-Reich (2020), if those spillover effects are symmetric, they are differenced out in the comparison between units, still allowing for correct identification of the local effects.[10]

[10]Moreover, Chodorow-Reich (2020) distinguishes between spillovers and endogenous reactions of "aggregate" treatments, such as interest policy. Here we include endogenous response of monetary policy as one potential source of spill-over.

The general framework laid out in equation (11) can also accommodate spillovers from a subset of local shocks by treating those as factors η_t to be included in the model. As such, the framework can control for confounding effects coming from such spillovers.

Heterogeneous Exposure Using our notation, Sarto (2024) (equation 3) writes his economic model as

$$y_{i,t} = C^i_{G,y} G_t + C^i_{g,y} \Delta g_{i,t} + \bar{B}^i_y \eta_t + \tilde{\epsilon}^i_{y,t},$$

where the coefficients of $\bar{B}^i_y$ are set such that the aggregate policy shocks η^g_t are excluded from the equation. Sarto (2024) differentiates between three types of elasticities: micro-local ($C^i_{g,y}$), micro-global ($C^i_{G,y}$) and macro, obtained from aggregating the equation above. With homogenous micro-global elasticity $C^i_{g,y}$, the micro-local elasticity corresponds to m^{local}, whereas the micro-global elasticity equals to m^{agg}. With heterogeneous global elasticity the relationship is given by

$$m^{i,local} = \frac{G^i_{G,y} - m^{agg}}{B^i_{\eta^g,g}} + C^i_{g,y},$$

That is, the local effect of η^G_t incorporates both the local, partial-equilibrium effect from the difference between the global and aggregate effects, normalized by the local effect of the aggregate shock, $B^i_{\eta^g,g}$. Note that the local effect then conflates the partial equilibrium effect of local shocks estimated by Wolf (2023) with the heterogeneous exposure to that shock encoded in $B^i_{\eta^g,g}$. Sarto (2024) is able to estimate the local elasticity $C^i_{g,y}$ separately from the global elasticity $C^i_{G,y}$ by adding structure to the local idiosyncratic shocks $\tilde{\epsilon}_{i,t}$.

4 Time Series Model

We now describe the full time-series model, which generalizes the model in Matthes and Schwartzman (2023). This allows one to handle more general environments then the one discussed so far. First, it encompasses dynamic environments by allowing for lagged dependent variables. More generally, one may be concerned about error in the elements of B_{η^g} obtained from microeconometric studies. We introduce a framework that incorporates persistence, heterogeneous responses to aggregate shocks across cross-sectional units, imperfect knowledge of the constraints on B^i implied by Assumptions P-1 and P-2.[11] It also safeguards against misspecification or estimation error in B_{η^g} by using existing 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 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.

As before, for each of I idiosyncratic units, we track K unit-specific variables. Those can be aggregated into an equal number of aggregate variables, to which we can add aggregate-only variables, for a total of $N \geq K$ 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

$$X_t^{agg} = \mu^{agg} + \sum_{l=1}^{L} A_l^{agg} X_{t-l}^{agg} + B^{agg}\eta_t + \varepsilon_t, \qquad (14)$$

where X_t^{agg} is an N dimensional vector collecting observed aggregate endogenous

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

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. μ^{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

$$X_t^i - X_t^{agg} = \mu^i + \sum_{l=1}^{L^{agg}} A_l^i X_{t-l}^{agg} + \sum_{l=1}^{L^{reg}} C_l^i X_{t-l}^i + B^i \eta_t + \varepsilon_t^i, \quad i = 1, \cdots, I \tag{15}$$

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. μ^i, A_l^i, C_l^i and B^i are conformable vectors and matrices of parameters. 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.

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 η_t or contemporaneous and lagged aggregate variables X_t^{agg}.

[a] In that case we simply need to modify the left-hand side of Equation (15) 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.

4.1 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

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

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

$$X_t = X_t^{agg} \otimes \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}_{N \times 1} + \mu^X + \sum_{l=1}^{L^{reg}} \tilde{A}_l^X X_{t-l} + \sum_{l=1}^{L^{agg}} \tilde{C}_l^{agg} X_{t-l}^{agg} + B^X \eta_t + \varepsilon_t^X \quad (16)$$

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:

$$Z_t = \mu^Z + \sum_{l=1}^{max(L^{agg},L^{reg},L)} A_l^Z Z_{t-l} + \underbrace{B^Z \eta_t + \varepsilon_t^Z}_{w_t^Z} \tag{17}$$

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 (14) into each idiosyncratic set of equations (15).

4.2 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. In particular, the prior for B^Z, which encodes our identification assumptions, is assumed to be Gaussian. 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.

When discussing the applications, 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. We detail the Gibbs sampler algorithm in appendix A and test the ability of the approach to identify the objects of interest under different conditions using a Monte Carlo exercise in appendix B.

Since we 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).

4.2.1 Prior on Aggregate Effect

The effect of η^g on the corresponding policy variables can be cast in terms of the proportion of its variance that it explains. We use this through the applications to set priors on that effect. In particular, for each policy variable $X^{agg}_{g,t}$ we suppose that a fraction θ of the variance of $X^{agg}_{g,t}$ is explained by the shock that we identify. Given that η^g_t has a unit variance, our prior mean for $B^{agg}_{\eta^g,g}$ is

$$E\left[B^{agg}_{\eta^g,g}\right] = (\theta\tilde{\Sigma}^{agg}_{g,g})^{1/2} \tag{18}$$

where $\tilde{\Sigma}^{agg}_{g,g}$ is the variance for the of the one-step ahead forecast error for policy variable obtained from estimating a version of our aggregate block via ordinary least squares (OLS).

4.2.2 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).[12] 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).

[12]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 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.

5.1 Data and Model Specification

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.

We estimate the model with a lag length of 2 for both the aggregate and state-level blocks and include $R = 3$ aggregate shocks in our estimation.

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 response of the aggregate variables to this shock is represented by the column of B^{agg} corresponding to the government spending shock η^g, $B^{agg}_{\eta^g} = (B^{agg}_{\eta^g,y}, B^{agg}_{\eta^g,g})'$ and the response of idiosyncratic variables is $B^i_{\eta^g} = (B^i_{\eta^g,y}, B^i_{\eta^g,g})'$ where, as before, $B^i_{\eta^g,y}$ is the reduced form effect of the shock η^g to variable y in unit i etc.

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

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 $B^i_{\eta^g,y}$ and $B^i_{\eta^g,g}$. To establish priors on the sensitivity of regional spending to the aggregate spending shock $B^i_{\eta^g,g}$, 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 β^i is used to inform the prior mean of $B^i_{\eta^g,g}$ after being rescaled by the effect of the government spending shock on aggregate government spending. Specifically, we set

$$E[B^i_{\eta^g,g}] = \beta^i E[B^{agg}_{\eta^g,g}],$$

where E here refers to the mean of the prior distribution for each parameter.

In the baseline specification described above, the 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 five years of the sample, a shift-share setup.[13]

[13]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

We further set the prior for the effect of the government spending shock on local output $B^i_{\eta^g,y}$ as

$$E[B^i_{\eta^g,y}] = \widehat{m}^{local-NS} E[B^i_{\eta^g,g}],$$

where $\widehat{m}^{local-NS}$ is the local multiplier estimated by Nakamura and Steinsson (2014)[14]

In both cases, the prior standard deviation for $B^i_{\eta^g,g}$ is set to half the absolute value of the prior mean. We again choose the prior standard deviation of $B^i_{\eta^g,y}$ 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 $B^i_{\eta^g,g}$ and $B^i_{\eta^g,y}$ average to zero for both specifications, consistent with our regional block specification in which we subtract X^{agg}_t from X^i_t.[15]

5.2.2 Priors on B^{agg}

To set the prior on $B^{agg}_{\eta^g,g}$ we follow the procedure delineated in Section 4.2.1 and choose the prior mean for $B^{agg}_{\eta^g,g}$ so that η^g accounts for a fraction θ of the variance of innovations to G, as estimated via OLS. We then choose θ to maximize the marginal likelihood of the estimated model.

To set the prior on the aggregate effect of the government shock on output we draw on the by now extensive literature on the topic. Specifically, we to establish the prior mean of $B^{agg}_{\eta^g,y}$ and the standard deviation of both $B^{agg}_{\eta^g,y}$, and $B^{agg}_{\eta^g,g}$ are implied by our prior for the government spending multiplier. To pin down these parameters, we draw $m^{agg} = B^{agg}_{\eta^g,y}/B^{agg}_{\eta^g,g}$ 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

statistic, so we find it useful to study both specifications here.

[14] $\widehat{m}^{local-NS} = 1.43$ for the baseline specification. The values for these multipliers are specification-specific and given in Appendix D.1.4.

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

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 (2018b): "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 C.[16]

5.3 The Aggregate Government Spending Multiplier

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]

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

[17]Although, as stressed by Nakamura and Steinsson (2018b), 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).

Table 1 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 with different prior-based identification schemes. For the case where we use the full identification scheme described above, 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.

The first column shows the prior distribution for m^{agg} derived from the literature. The prior mean is 0.8. With 68% probability the multiplier is between 0.5 1nd 1.18, and with 90% between 0.38 and 1.55. The prior distribution puts a probability of 28% for the the multiplier being above 1.

The second column shows what we obtain if we estimate the model without applying the priors to the local effects. In that case, as in the Cholesky benchmark above, the data adds virtually no new information to the prior, and the posterior remains unchanged.

The third column examines what happens if we add information on the sensitivity of local to aggregate government spending used to construct the instrument. This information proves relevant, increasing the posterior mean by 5 percentage points, and the probability of multiplier being above one to 34%.

The last column (our baseline) applies priors also to the local multipliers. This information is again relevant, raising the posterior mean by an additional 6 percentage points (thus 11 percentage points above the prior). It also significantly reduces the span of the 90% probability range for the multiplier from 1.65-0.4=1.25 to 1.25-0.56=0.69. The probability of the multiplier being larger than 1 raises further, to 0.37.

	(1) Prior	(2) Posterior	(3) Posterior	(4) 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
θ		1.00	0.25	1.00
Informative B_y^i prior		No	No	Yes
Informative B_g^i prior		No	Yes	Yes

Table 1: 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 2. 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 1)[19]

[19]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.

	(1) Posterior	(2) Posterior
	0.90	0.97
m^{agg}	(0.41, 1.71)	(0.49, 1.52)
	[0.59, 1.32]	[0.68, 1.29]
$Prob(m^{agg} > 1)$	0.39	0.46
Log MDD	-7208.78	-7292.27
θ	1.00	1.00
Informative B_y^i	No	Yes
Informative B_g^i	Yes	Yes

Table 2: 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 Robustness

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 D.

First, we investigate how sensitive the results are to alternative data transformation. We use output-weighted averages of regional data for aggregates rather than equally weighted table (A-2). Results are virtually unchanged. We also take 1-year differences rather than 2 for the data, shortening the horizon for the multiplier. We find that, whereas, the posterior mean is now multiplier is smaller, sticking closer to the prior mean of 0.8, the use of cross-sectional data is informative in that it consistently tightens the interval around the posterior mean (table A-3).

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 (table A-4), we add lagged dependent variables to the first-stage regression used to estimate the effect of federal spending on local spending, so as to make it consistent with the main regression specification. Interestingly, allowing for controls leads to a larger update of the multiplier towards one.

We further investigate the role played by loosening the aggregate prior (tables A-6 and A-7). We find the estimated path for η_t^g remains highly correlated with the original estimate (see table A-9), but the posterior bands around the aggregate multipliers increase significantly. This confirms that, while the estimated shock is robustly estimated, the information content of the estimated shock for the estimation of the aggregate multiplier is relatively modest given the short sample and annual data.

Last, we investigate the role of two features of our benchmark prior: the standard deviation of the priors on local effects and the choice of θ (the fraction of the variance of the government spending accounted for by the shock). Table A-8 shows how the estimates change if we bring the prior uncertainty on local effects to zero, effectively asserting perfect certainty about the local effects. This brings the posterior estimate of the aggregate multiplier very close to 1. Figure A-5 shows the aggregate multiplier for different degrees of prior precision for those local effects. Not surprisingly, the aggregate multiplier declines as we become less certain about the local effects. The estimated multiplier is similarly sensitive to the choice of θ (figure A-6), but most of the changes occur for values of θ associated with posterior density significantly below the benchmark.

The first row of table A-9 shows that the sectoral priors add substantial information about the path of the shock, since the estimated shock is otherwise only weakly correlated. Otherwise, under all alternative specifications retaining the two-year horizon, the estimated path for η_t^g remains highly correlated with the original estimate, indicating that the method can robustly estimate η_t^g.

6 Application # 2: Chodorow-Reich, Nenov, and Simsek (2021)

As a second application, we revisit Chodorow-Reich et al. (2021). Their focus is in the effect of a change in stock market wealth to the real economy, in particular, the effect to the local labor market (payroll and employment). They interpret their results through a regional economy model economy where news about future productivity affects stock market wealth and, through it, the consumption of stock holders.

6.1 Empirical Specification and Data

The main empirical specification in the paper is

$$\Delta^i_{t-1,t+h} y = \beta_h s^i_{t-1} R^i_{t-1,t} + \Gamma'_h X^i_{t-1} + \varepsilon^i_{t-1,t+h} \tag{19}$$

for county i and quarter t at horizon h, where $\Delta^i_{t-1,t+h} y$ is the change in either local employment or payroll from $t-1$ to $t+h$, $s^i_{t-1} = S^i_{t-1}/W^i_{t-1}$ is the measure of local stock market weath $S_{i,t-1}$ normalized by local wage W^i_{t-1}, and $R^i_{t-1,t}$ is the return on the area-specific stock portfolio. The return on the local portfolio is calculated as

$$\begin{aligned} R^i_{t-1,t} &= b^i_t R^m_{t-1,t} + (1 - b^i_t) R^f_{t-1,t} \\ &\approx b^i_t R^m_{t-1,t} \end{aligned} \tag{20}$$

where $R^m_{t-1,t}$ is the stock market return, $R^f_{t-1,t}$ is the risk-free rate, and $b_{i,t}$ is the county specific beta, tying local portfolio return to the return on the stock market. The specification is estimated by local projection using OLS.

To construct $b_{i,t}$ the authors use the relationship between market beta and age from Barber and Odean (2000), and the county age-wealth distribution. The specification also includes a vector of controls X^i_{t-1}, including county fixed effect, state $\times$ quarter fixed effect, eight lags of $s^i_{t-1} R^i_{t-1,t}$, and interactions between $s_{i,t-1}$ and holding return on a 5-year Treasury bond, growth of national house prices from $t-1$ to t, $s^i_{t-1}\times$ growth of national labor income and noncorporate business income from $t-1$ to the cumulative total over next 12 quarters, predicted employment growth based on industry composition (Bartik shift-share measure). We refer the reader to their paper for further details.

We are interested in inferring the impact of stock market wealth on national employment and wage bill. Therefore, the aggregate variables X_t^{agg} include quarterly growth of the aggregate employment and payroll, the aggregate $s_{t-1}^{agg}R_{t-1,t}^{agg}$ and the controls using interactions between stock holdings s_{t-1}^{agg} and the 5-year Treasury bond, a national index of house prices and the growth to national labor income and noncorporate business income.[20] The regional variables X_t^i include the regional counterpart of the aggregate variables listed above (employment and wage bill are measured at the county-level at a quarterly frequency using QCEW data), and predicted employment growth based on industry composition (Bartik shift-share measure).[21]

6.2 Identification via Priors

As before, for standard VAR-type parameters, we use Minnesota priors Doan et al. (1984), as is common in the literature. Except for the priors in the elements of B^Z described below, the priors are summarized in table A-1.

We denote the element of η representing the stock-market wealth shock by η^r. This can be viewed, as in Chodorow-Reich et al. (2021), as stemming from stock market fluctuations that are disconnected from current labor market conditions, such as news and uncertainty shocks in environments without wealth effects on labor supply, or non-fundamental shock to asset pricing due to bubbles or changing liquidity conditions. Hence, the response of the aggregate wage bill, employment and stock market wealth are, respectively, represented by the first column of B^{agg}, which we call $B_{\eta^r}^{agg} = (B_{\eta^r,\ell}^{agg}, B_{\eta^r,w\ell}^{agg}, B_{\eta^r,sR}^{agg})'$ and the response of idiosyncratic variables is

[20] We calculate aggregate $s_{t-1}^{agg}R_{t-1,t}^{agg}$ by taking a wage-bill aggregate of local $s_{t-1}^{i}R_{t-1,t}^{i}$, so that $s_{t-1}^{agg}R_{t-1,t}^{agg} = \sum_j \frac{W_{t-1}^j}{W_{t-1}^{agg}} \frac{S_{t-1}^j b_t^j}{W_{t-1}^j} R_{t-1,t}^m = \sum_j \frac{W_{t-1}^j}{W_{t-1}^{agg}} \frac{S_{t-1}^j b_t^j}{W_{t-1}^j} R_{t-1,t}^m$

[21] When constructing $X_t^i - X_t^{agg}$ for the Bartik shift-share method, we subtract the national employment growth.

$B^i_{\eta^r} = (B^i_{\eta^r,\ell}, B^i_{\eta^r,w\ell}, B^i_{\eta^r,sR})'$, where again for legibility we index row elements of each B matrix by the variable letter rather than its position in the VAR.

6.2.1 Regional Priors

Prior on $\mathbf{B^i_{\eta^r,sR}}$. We put the informative prior restriction on the element of η_t corresponding to the first element in the equation involving $s^i_{t-1}R^i_{t-1,t} - s^{agg}_{t-1}R^{agg}_{t-1,t}$. Recall that s^i_{t-1} is defined to be S^i_{t-1}/W^i_{t-1}, the stock market wealth normalized by wage.

If we interpret η^r_t as a shock to stock market return $R^m_{t-1,t}$, then its impact on region i wealth/income ratio will be in proportion to $s^i_{t-1}b^i_{t-1}$, the ratio of stock market wealth to income multiplied by the region b^i_{t-1}. Since this varies little over time[22], we set $E[B^i_{\eta^r,sR}]$ to be proportional to the average value for $s^i_{t-1}b^i_{t-1} - s^{agg}_{t-1}b^{agg}_{t-1}$ over the first 5 years, with $b^{agg}_t = \sum_j S^j_t b^j_t / S^{agg}_t$ the national beta of stock market portfolios on stock returns. In particular, we set

$$E[B^i_{\eta^r,sR}] = \frac{s^i_{t-1}b^i_{t-1} - s^{agg}_{t-1}b^{agg}_{t-1}}{s^{agg}_{t-1}b^{agg}_{t-1}} E[B^{agg}_{\eta^r,sR}], \tag{21}$$

where the denominator adjusts for the fact that the impact of a shock to stock market returns on aggregate wealth is proportional to $s^{agg}_{t-1}b^{agg}_{t-1}$. We take the time average for the first five years of the sample of $\frac{s^i_{t-1}b^i_{t-1} - s^{agg}_{t-1}b^{agg}_{t-1}}{s^{agg}_{t-1}b^{agg}_{t-1}}$ to settle on the prior mean. This follows the standard practice of shift-share instruments to mitigate endogeneity concerns.

Prior on $\mathbf{B^i_{\eta^r,\ell}}$ and $\mathbf{B^i_{\eta^r,w\ell}}$. From (19) with $h = 0$ estimated using the replication package made available by Chodorow-Reich et al. (2021), we obtain $\hat{\beta}_0$ for log employment and wage bill.[23]. The prior mean of $B^i_{\eta^r,\ell}$ in employment equation and $B^i_{w\ell,1}$ in wage bill equation are given by the corresponding $\hat{\beta}_0$ multiplied by the prior

[22] We compute absolute values of coefficients of variation for the coefficients of equation (21) for each county i. The median is 0.30

[23] Regional β_0 is -0.103 for employment and 0.411 for wage bills

mean of $B^i_{\eta^r,sR}$ described above. The prior standard deviation is again half of absolute value of prior mean.

Prior on Other Elements of First Column of $\mathbf{B^i}$. We place the prior mean of zero, and the prior standard deviation is 2.5 times prior mean of $B^i_{\eta^r,sR}$.

6.2.2 Aggregate Prior

For all elements of B^{agg}, the prior standard deviation takes large value, 10.0. The prior mean is zero except for $B^{agg}_{\eta^r,sR}$, contemporaneous response of stock market wealth to an identified shock. Following the strategy in the government multiplier exercise, we first estimate a VAR with only Y^{agg}_t to obtain the variance-covariance matrix $\tilde{\Sigma}^{agg}$. The prior mean of $B^{agg}_{\eta^r,sR}$ is chosen to be $\theta(\tilde{\Sigma}^{agg}_{sR,sR})^{1/2}$ where θ is a constant between 0 and 1. We report here estimation results for $\theta = 0.25$.[24]

6.3 Response of Employment and Wage Bills

To be consistent with the object of interest in the original paper, we report

$$\beta_h = \frac{\text{Cumulative IRF of Employment or Wage at Horizon } h}{\text{IRF of Stock Market Wealth at Impact}}$$

Given that they are using log difference of employment and wage, this can be interpreted as “a percentage increase in employment or wage at horizon h driven by a one-dollar increase in stock market wealth in response to a shock to η_t at impact”.

The results are presented in figure 1. The wealth shock generates large, significant, and sustained increases in employment and wage bill. The initial impact is on the wage bill, but employment catches up eventually.

Based on the multi-region structural model, Chodorow-Reich et al. (2021) calibrate lower bounds of percentage changes in aggregate employment and wage when there is no monetary policy response. In the absence of the monetary policy reaction, the

[24] Calculating the MDD in this instance is very costly in terms of memory since, compared to NS, we include more variables and there are over 2,500 regions (There were only 51 in NS).

aggregate employment increases at least 1.3 percent after two years, and the aggregate wage increases at least 3.23 percent after two years. The posterior median of the response from the VAR at $h = 7$ is 2.67 percent for employment and 1.67 percent for the wage bill. The wage bill growth is below the value implied by the theory, implying that the employment reduces due to the monetary policy in response to the shock to stock return. The response of employment to the stock return shock, however is so strong that it remains above the theory-implied value even in the presence of monetary policy response. Lastly, note that the posterior mean for the growth in wages per worker is in fact negative, as the wage bill grows below employment. However, probability ranges around the posterior mean are wide enough that the change in wages is not significantly different from zero.

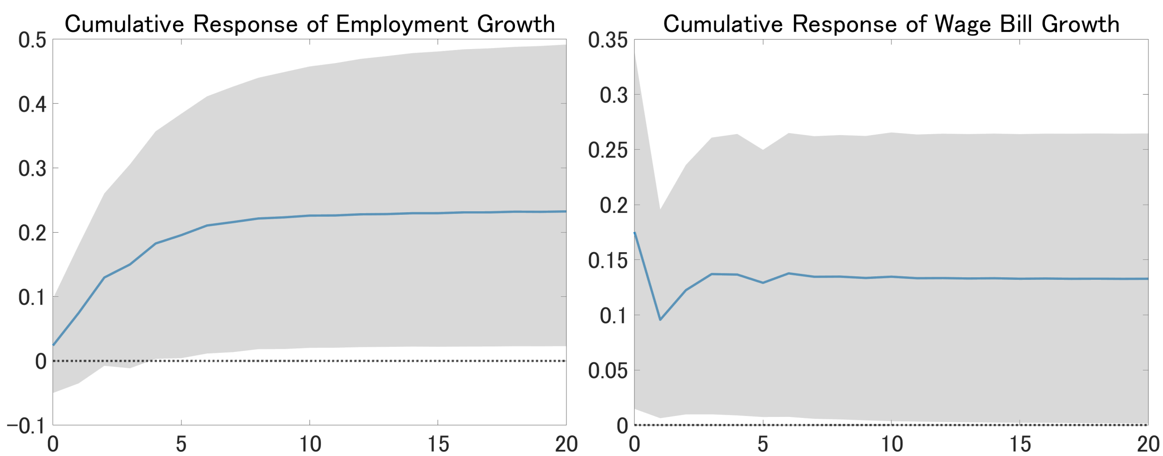

Figure 1: Cumulative responses of employment and wage bill to stock market wealth. Grey bands correspond to 90% probability ranges.

	Chodorow-Reich et al. (2021) lower bound	Posterior median estimates
$\beta_7^{\text{employment}}$	1.29	2.67 [1.22, 4.24]
$\beta_7^{\text{wage bill}}$	3.23	1.67 [0.74,2.68]
$\beta_7^{\text{wage bill}} - \beta_7^{\text{employment}}$	—	-0.98 [-2.62,0.58]

Table 3: β_7 for Wage Bill Growth

Table 4: Aggregate effects of stock market wealth. The Chodorow-Reich et al. (2021) lower bound is calculated based on the assumption that there is no monetary policy response to the stock market shock. Values in square brackets denote 68% probability range for the posterior estimates.

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.

Barber, B. M. and T. Odean (2000). Trading is hazardous to your wealth: The common stock investment performance of individual investors. *The journal of Finance 55*(2), 773–806.

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 (2018). Inference in structural vector autoregressions when the identifying assumptions are not fully believed: Re-evaluating the role of monetary policy in economic fluctuations. *Journal of Monetary Economics 100*(C), 48–65.

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*.

Chodorow-Reich, G., P. T. Nenov, and A. Simsek (2021). Stock market wealth and the real economy: A local labor market approach. *American Economic Review 111*(5), 1613–1657.

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.

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 (2018a). Identification in macroeconomics. *Journal of Economic Perspectives 32*(3), 59–86.

Nakamura, E. and J. Steinsson (2018b, 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. (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 The Gibbs Sampler Algorithm

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 (μ^{agg}, $\{A_l^{agg}\}_{l=1}^{L}$, B^{agg}, Σ^{agg}) and the regional block (μ^i, $\{A_l^i\}_{l=1}^{L^{agg}}$, $\{C_l^i\}_{l=1}^{L^{reg}}$, B^i, Σ^i $\forall i = 1, ..., N$) η_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 (μ^{agg}, $\{A_l^{agg}\}_{l=1}^{L}$, B^{agg}, Σ^{agg}) conditional on regional variables and η_t can be drawn using known conditional distributions (we assume Gaussian priors for B^{agg}).

- Regional variables (μ^i, $\{A_l^i\}_{l=1}^{L^{agg}}$, $\{C_l^i\}_{l=1}^{L^{reg}}$, B^i, Σ^i $\forall i = 1, ..., N$) conditional on aggregate variables and η_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.

B 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 for the government spending application (reported in Table 1) as the data generating process (DGP)[1]. The prior of $B_{g,1}^{agg}$ and $B_{y,1}^{agg}$ are largely uninformative -

[1]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

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.[2] 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.

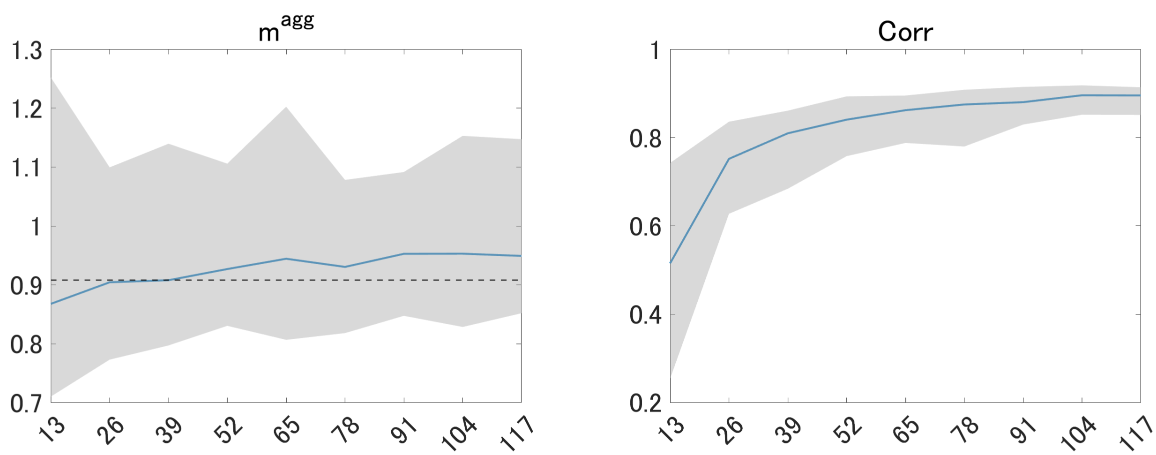

Figure A-1: Sensitivity to the number of time series observations T ($N = 51$)

---

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 36 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).

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

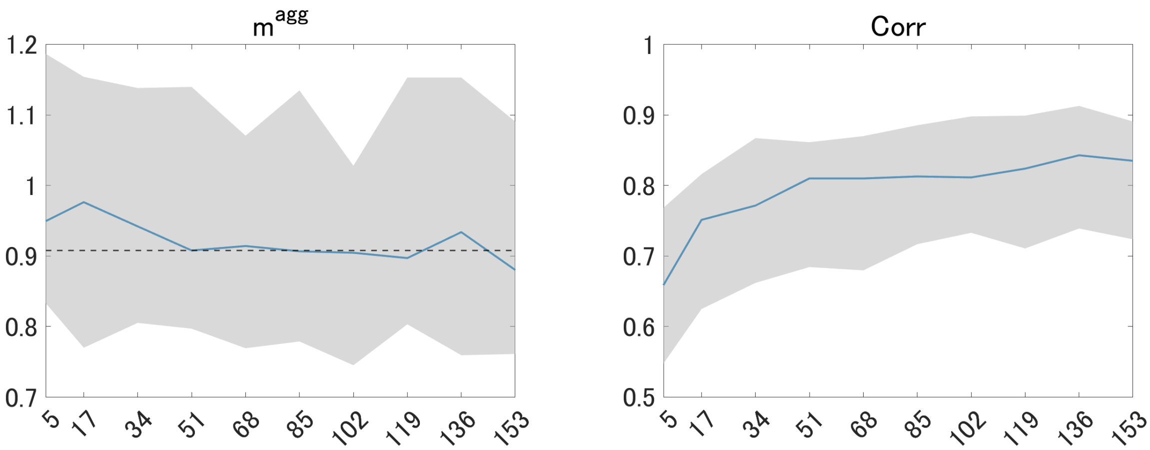

Figure A-2: Sensitivity to the number of cross-sectional units N ($T = 39$).

Figures A-1 and A-2 explore the sensitivity to the sample length and the size of the cross-section respectively. The solid blue line of the left panels 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 right panels report 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 only slight 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.

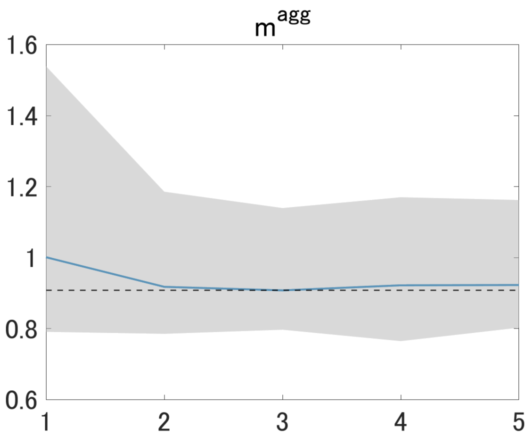

Figure A-3: Sensitivity to R ($R^{true} = 3$)

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 A-3 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 A-4), 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.

C 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 (14) and (15) without acknowledging the factor structure in the forecast errors and set the estimated $\tilde{\Sigma}^{agg}$ and $\tilde{\Sigma}^i$ $(i = 1, \cdots, N)$

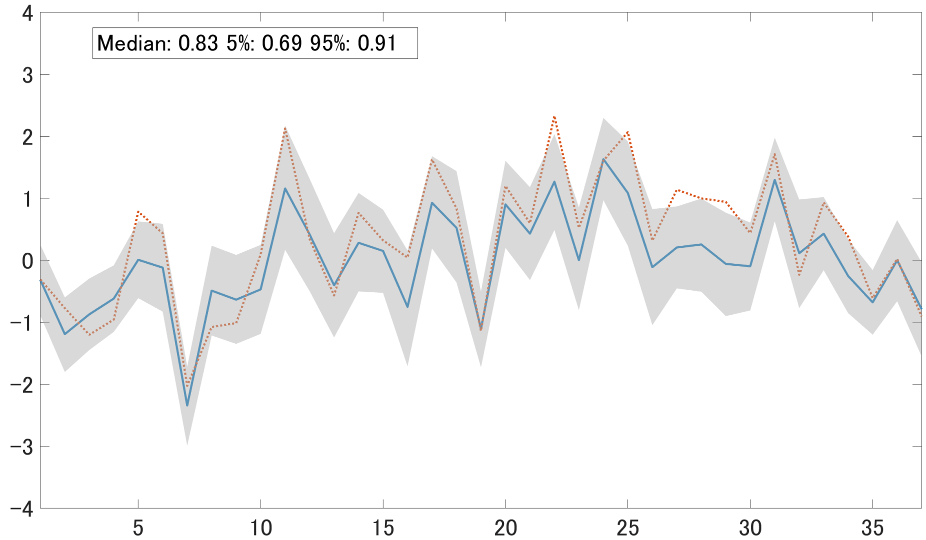

Figure A-4: One simulated shock series along with the estimated shock and 90 percent posterior bands for that sample. $(T, N) = (39, 51)$. Legend gives percentiles of the distribution of correlation between the true shock and estimated shock.

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 θ (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.

C.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 σ_i and σ_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 (14) and (15) without factors. The prior hyperparameters are set as $\phi_0 = 0.2$, $\phi_1 = 0.5$, $\phi_2 = 10^5$, and $h(l) = l$.

For the government spending application, we adopt the following priors for the variables not discussed in the text:

	Type of Distribution	Parameters
Aggregate Block		
μ^{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
Σ^{agg}	Inverse Wishart	Scale: OLS dof: 10
Regional Block		
μ^{i}, C^{i}	Normal	Minnesota Prior
A^{i}	Normal	Mean: 0.0, Std: 0.5
B^{i} (Identified)	Normal	Regional information (See main text)
B^{i} (Unidentified)	Normal	Mean: 0.0, Std: 10
Σ^{i}	Inverse Wishart	Scale: OLS dof: 10

Table A-1: Prior Specifications for Aggregate and Regional Blocks

D More Results for Government Spending Application

D.1 Transformed Data

D.1.1 Output-Weighted Aggregate Data

We compute the weights of output in each state relative to the aggregate output, and take time average of them. We construct aggregate variables by taking the weighted average of regional variables using the averaged output weights. We estimate the model with the alternative aggregate variables, where we find very similar posterior to the baseline.

	(1) Prior	(2) Posterior (First-Stage)	(3) 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
θ		1.00	1.00
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.

D.1.2 One-year differences

We estimate the model with the one-year difference of government spending and output.[3] Specifically, we let

$$y_t^{agg} = \frac{Y_t^{agg} - Y_{t-1}^{agg}}{Y_{t-1}^{agg}}, \quad g_t^{agg} = \frac{G_t^{agg} - G_{t-1}^{agg}}{Y_{t-1}^{agg}}, \quad y_t^i = \frac{Y_t^i - Y_{t-1}^i}{Y_{t-1}^i}, \quad g_t^i = \frac{G_t^i - G_{t-1}^i}{Y_{t-1}^i}$$

We follow the same strategy for choosing the prior as our baseline estimation, while the prior for regional parameters is adjusted accordingly by re-estimating the Nakamura and Steinsson (2014) regression with the alternative data. The one-year aggregate multiplier is smaller than the two-year multiplier, while our estimate is still in line with other evidence on the multiplier.

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

	(1) Prior	(2) Posterior (First-Stage)	(3) 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
θ		0.50	0.65

Table A-3: Observables based on one-year differences. 90% posterior bands are in parentheses, and 68% bands are in square brackets.

D.1.3 Alternative First-Stage Regression

We estimate the alternative IV regression for Nakamura and Steinsson (2014) by including lags of output and government spending.

$$y_t^i = m^{local} g_t^i + \sum_{l=1}^{2} \phi_{g,l} g_{t-l}^i + \sum_{l=1}^{2} \phi_{y,l} y_{t-l}^i + \alpha_i + \gamma_t + \varepsilon_t^i$$
$$g_t^i = \beta^i g_t^{agg} + \sum_{l=1}^{2} \phi_{g,l} g_{t-l}^i + \sum_{l=1}^{2} \phi_{y,l} y_{t-l}^i + \alpha_i + \gamma_t + \varepsilon_t^i$$

Re-estimating our VAR model with the alternative β^i and m^{local}, we get the posterior multiplier consistent with our baseline.

	(1) Prior	(2) Posterior	(3) Posterior
m^{agg}	0.80	0.87	0.96
	(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
θ		0.30	1.00
Informative B_y^i		No	Yes

Table A-4: First-stage regression now includes same controls as our baseline model. 90% posterior bands are in parentheses, and 68% bands are in square brackets.

D.1.4 Local multipliers from different data transformations

We report below the local multipliers from different data transformations:

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

Table A-5: Local multiplier estimates used to set priors, obtained by estimating the cross-sectional regression under different data transformations. See text for details.

D.2 Different Choice of Priors

D.2.1 Looser Aggregate Prior

	(1) Prior	(2) Posterior	(3) 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
θ		0.25	1.00
Informative B_y^i		No	Yes

Table A-6: 90% posterior bands are in parentheses, and 68% bands are in square brackets.

We change the prior distribution for aggregate multiplier to be looser than the baseline. Now, the 90% prior interval includes 0 and 3. The posterior median of the multiplier is slightly above the baseline estimate, while it comes with a wider posterior interval.

D.2.2 Even Looser Aggregate Prior

	(1) Prior	(2) Posterior	(3) 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
θ		0.30	1.00
Informative B_y^i		No	Yes

Table A-7: 90% posterior bands are in parentheses, and 68% bands are in square brackets.

We loosen the aggregate prior further, covering ± 6.3 as the 90% interval. The estimates are less sensible to others, suggesting the role of aggregate prior in relatively short time series like Nakamura and Steinsson (2014) data.

D.2.3 Standard deviation of local priors

We evaluate 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.

We first think about the situation where an econometrician is very sure about the regional coefficients. We estimate the model with the identical prior setting as in the last column of Table 1, except that the prior standard deviations for $B^i_{g,1}$ and $B^i_{y,1}$ are very small: 10^{-7}.

	(1) Prior	(2) Posterior
	0.800	0.978
m^{agg}	(0.376, 1.548)	(0.502, 1.509)
	[0.528, 1.177]	[0.686, 1.288]
$Prob(m^{agg} > 1)$	0.282	0.470
Log MDD		-7270.9465
θ		1.0
Informative B_y^i prior		Yes
Informative B_g^i prior		Yes

Table A-8: Results based on Perfectly Precise Information on Local Coefficients

Figure A-5 shows what happens when we use other values than 0.5. Not surprisingly, the less confident one is about the local effects, the more the aggregate multiplier estimate converges towards the prior. Interestingly, the log likelihood is also largest for tight priors.

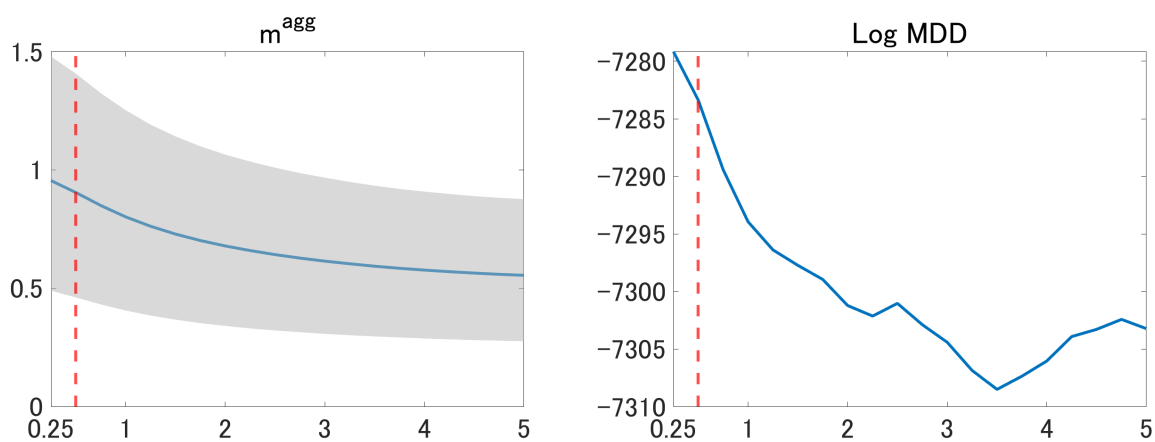

Figure A-5: Changing standard deviation of prior on local effects. Left panel plots the aggregate multiplier (median and 90 percent posterior bands), right panel plots the marginal data density estimated via method in Geweke (1999). Dashed red vertical line shows the benchmark value.

D.2.4 Choice of θ

We have thus far picked θ to maximize the marginal likelihood. How much does this matter? Figure A-6 gives an answer.

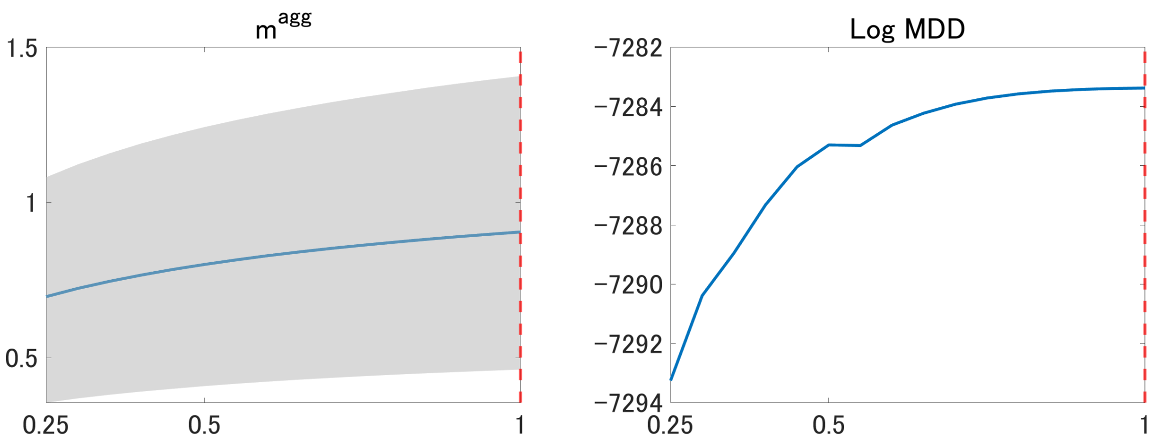

Figure A-6: Changing θ. Left panel plots the aggregate multiplier (median and 90 percent posterior bands), right panel plots the marginal data density estimated via method in Geweke (1999). Dashed red vertical line shows the benchmark θ value.

The fit of the model increases substantially with $\theta = 1$, as does the posterior estimate for the aggregate multiplier.

D.3 Correlation of Identified Shocks Across Specifications

Uninformative B_y^i and B_g^i (Table 1 Column 2)	0.1297
Output weighted mean as aggregates (Table A-2 Column 2)	0.9988
Alternative first-stage (Table A-4 Column 2)	0.9962
Loose aggregate prior (Table A-6 Column 3)	1.0000
Even looser aggregate prior (Table A-7 Column 3)	0.9998
Certain local information (Table A-8 Column 2)	0.9934

Table A-9: Correlation of posterior medians of estimated shocks with baseline

To see how robustly we identify the shock of interest, we compute the correlation between the posterior median of identified shocks in our baseline estimation (Table 1 Column 4) with the one from alternative specifications. To make a fair comparison, we do this for the specifications where we use two-year difference, the NS first-stage prior, and informative B_y^i. We find very strong correlation across different specifications, suggesting that the identified shock is almost identical.

E Local Multiplier With 90 Percent Bands

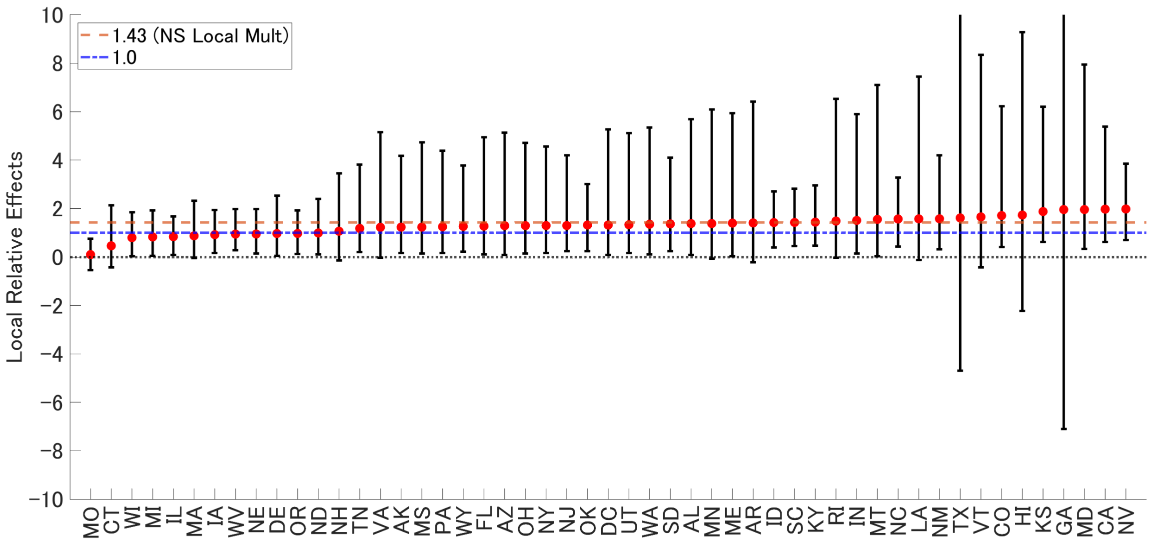

Figure A-7: Median of Local Relative Multiplier with 90% Posterior Interval