Chapter 8

Empirical strategies and quantitative macroeconomics

8.1 Introduction

We often have questions that require quantitative answers. Can a certain theory explain the amount of growth or volatility we see in the data? How do the benefits of a proposed policy compare to the costs? Quantitative answers require empirical evidence. Perhaps there is evidence that directly speaks to the question of interest. More often there is not and we need to bridge the gap between the available evidence and the answer to our question with theoretical models or assumptions.

The types of questions we contemplate typically require estimates of causal effects that measure the consequences of changing one aspect of the economy holding the other economic fundamentals constant. A key challenge for economists, and macroeconomists in particular, is that it is often impossible to perform experiments on the economy. Instead, we have to estimate these causal effects using data that reflect equilibrium outcomes in which many variables are jointly determined and responding to a variety of different shocks. In this chapter we will discuss four different strategies to address this challenge.

We start with two strategies that impose relatively fewer assumptions on the analysis starting with **natural experiments**. While we usually cannot perform experiments on the economy it may be the case that there is variation in the data that is effectively exogenous with respect to the other variables of interest. If we can find such natural experiments, then we can estimate causal relationships more or less directly. We then turn to techniques for identifying causal effects from time series data. A vector autoregression (VAR) is a statistical tool that expresses the dynamic evolution of a group of time series in terms of the lags of the same time series. In a **structural VAR**, we add additional identifying assumptions to a VAR to uncover patterns in the data that can be given a causal interpretation.

The last two strategies we consider rely closely on economic theory in the form of a fully-specified model. A model can be used as a laboratory in which we can perform experiments to measure causal effects and experiment with proposed policies. The key issue then is how we use data to inform the model and its parameters. **Structural estimation** treats the model as the data-generating process and selects parameters to best explain the observed data. The strategy of **calibration** considers each parameter of the model and seeks evidence that speaks to the value of that parameter. When we calibrate a model, we do not presume that the model explains all of the variation in the data we observe.

The rest of this chapter will demonstrate these four different strategies and how they would answer the question "how does GDP respond to an increase in government spending?" The answer to this question is often summarized by a single quantity known as the "fiscal multiplier." In section 8.2, we begin by setting out the core issue we face in identifying an exogenous change in government spending in order to estimate the fiscal multiplier. We then turn to the empirical strategies.

8.2 The identification challenge

The "fiscal multiplier" refers to the *causal* effect of an increase in government spending (G_t) on output (y_t). By causal effect we mean G_t increases independently of other forces affecting the economy. Suppose we have time series data on G_t and y_t. Then we could simply look at how strongly the two co-move. However, to estimate a causal effect rather than a correlation we need the increase in spending to be exogenous with respect to other variables of interest. For example, it could be the case that an increase in productivity creates more resources and the economy chooses to spend more on public goods. In that case, the increase in output is not caused by the increase in spending.

In order to illustrate the identification challenge, we will consider a simple static model of the economy. The economy is populated by a representative household with preferences

$$U_t = \log c_t - \frac{\ell_t^{1+\psi}}{1+\psi} + \gamma \eta_t \log G_t, \tag{8.1}$$

where c_t is consumption, ℓ_t is labor supply, and G_t is spending on public goods. The parameter γ determines the steady state preference for public goods while η_t is an exogenous fiscal shock that shifts the value of public goods over time. For instance, government purchases may be especially useful in a time of war, which would correspond to an increase in η_t. The economy produces goods out of labor according to $y_t = A_t \ell_t$ where A_t is an exogenous TFP level. Goods can be used for consumption or public goods so we have $y_t = c_t + G_t$. Suppose a planner chooses $\{c_t, \ell_t, G_t\}$ to maximize the household's utility subject to $c_t + G_t = A_t \ell_t$. The solution to this problem results in the first-order conditions

$$\frac{A_t}{c_t} = \ell_t^{\psi}$$

and

$$G_t = \gamma \eta_t c_t.$$

Using the aggregate resource constraint to eliminate $c_t = y_t - G_t$ and the production function to eliminate $\ell_t = y_t / A_t$ we obtain

$$y_t = A_t \left(1 - \frac{G_t}{y_t}\right)^{-\frac{1}{1+\psi}} \tag{8.2}$$

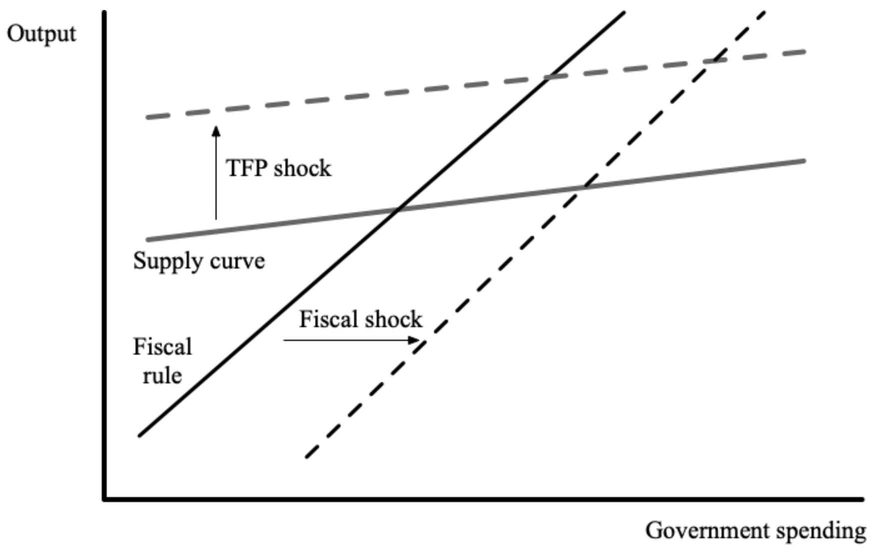

Figure 8.1: Equilibrium of supply curve (8.2) and fiscal rule (8.3).

and

$$G_t = \frac{\gamma\eta_t}{1+\gamma\eta_t} y_t. \tag{8.3}$$

Together, these two conditions define the equilibrium in (G, y) space, shown in Figure 8.1. Equation (8.2) reflects the supply side of the economy. This equation depends on G/y due to wealth effects on labor supply. As more resources are devoted to the government and fewer to consumption, the marginal utility of consumption rises relative to the marginal disutility of work leading the planner to increase labor supply. Equation (8.3) is the fiscal policy rule—G_t increases if y_t or η_t increase. An increase in η_t clearly raises the benefit of public spending. An increase in y_t lowers the opportunity cost of allocating resources to G_t rather than c_t.

The fiscal multiplier is the change in output (in levels) divided by the change in spending (in levels). A fiscal multiplier equal to one means that one dollar more of spending raises GDP by one dollar. Crucially, the change in spending must be exogenous because the fiscal multiplier is meant to be a causal effect—it answers the question "what would happen if G increases *ceteris paribus*?" One way to calculate the fiscal multiplier in a model is to replace the spending policy rule with one in which G is simply exogenous. One would then consider the equilibrium response of output to exogenous changes in G. Graphically, the fiscal multiplier is the slope of the supply curve in Figure 8.1. An alternative approach to calculating the fiscal multiplier is to find an exogenous force that shifts the fiscal policy rule. In our model, the fiscal shock η_t is an exogenous shift of the fiscal policy rule. We can then calculate the fiscal multiplier as $(dy_t/d\eta_t)/(dG_t/d\eta_t)$. Or, in words, the change in y induced by a change in η scaled by the change in G that was induced by η.

Now suppose we have data on G_t and y_t over time. These data reflect movements in A_t as well as η_t. Changes in A_t shift the supply curve while changes in η_t shift the fiscal rule. As both of the curves in Figure 8.1 are subject to unobserved shocks, the empirical relationship between y_t and G_t does not reveal the slope of either curve. In order to calculate the fiscal multiplier, we need to isolate shifts in the fiscal rule that move the economy along the supply curve. The empirical strategies that follow all must confront this identification challenge.

Defining the fiscal multiplier in a dynamic economy

In a dynamic economy, a change in spending can be more or less persistent and the persistence of spending will matter for how the economy reacts. Moreover, an increase in spending at one date will lead to a change in output on impact and at subsequent dates. In this dynamic setting, the fiscal multiplier is not a single number, but depends on the details of the expected path of G_t and involves an impulse response of y_t.[a] The literature has used several methods to reduce this complexity to one dimension. Blanchard and Perotti (2002) define the multiplier as the ratio of the peak output response relative to the peak spending response. Alternatively, others follow Mountford and Uhlig (2009) in calculating the cumulative multiplier that integrates the impulse response of GDP and divides by the integral of the change in spending.

[a]Moreover, the response of y_t may depend on the state of the economy in which we start the experiment although in this chapter we will focus on first-order approximations to the dynamics of the economy that do not capture this state dependence.

8.3 Natural experiments

To assess the fiscal multiplier, we need a shift in fiscal policy that is unrelated to the state of the economy. Some of the largest changes in fiscal policy have occurred at times of war. It is reasonable to argue that these wars were not caused by economic conditions and so changes in government spending due to wars can be treated as exogenous. This is an example of using a natural experiment for identification. Given our knowledge of the economy, we may be able to find an exogenous variable (e.g. war spending) that is correlated with the endogenous variable of interest (e.g. total government spending). In our system (8.2)–(8.3), we seek a variable z_t that is correlated with G_t but uncorrelated with A_t.

A series of papers has used exactly this strategy of studying military spending to measure the fiscal multiplier. One challenge this strategy has to overcome is that the change in spending may have been known ahead of time. If wealth effects on consumption and labor supply are important mechanisms in the transmission of fiscal shocks to the rest of the economy, these effects should occur when the agents learn about the change in fiscal policy rather than when the spending actually occurs. For example, it may be clear that war is on the horizon before the government increases spending. Therefore the economic impact of the war spending may be felt before the spending occurs. To overcome this challenge, Ramey (2011) measures the change in the expected present value of defense spending each quarter by reading historical newspaper articles to understand the timing of information about military spending. This data is shown in Figure 8.2 where we see that there were large changes in expected military spending in WWII and the Korean War. Ramey estimates the impulse responses of government spending and GDP to this news and uses those impulse responses to estimate the fiscal multiplier. She finds a multiplier near 1 when the sample period includes WWII and multipliers near 0.7 when WWII is excluded.

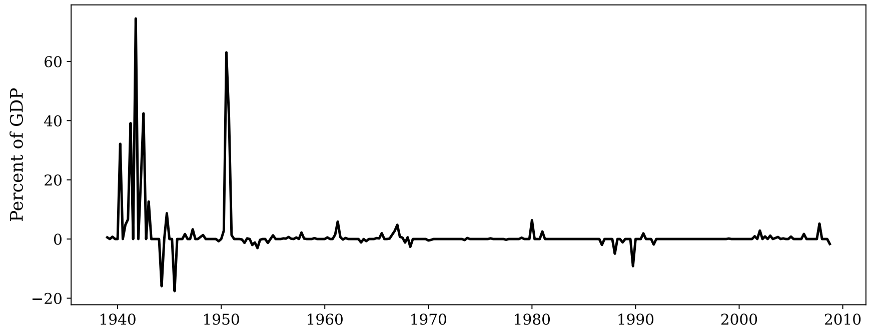

Figure 8.2: Change in the expected present value of military spending computed by Ramey (2011).

From natural experiments to impulse response functions: local projection

Let z_t be the measure of military spending news. As z_t is exogenous, we could simply regress output y_t on z_t. But perhaps there is a delay between the defense spending news and the output response. We could then regress y_{t+h} on z_t to estimate how output responds to spending news h periods after the news arrives. By varying h we can trace out a whole impulse response function. This procedure is called a local projection or Jordà (2005) projection.

Analyzing a natural experiment often requires using unconventional data sources. For example, Ramey used historical newspaper articles to construct her dataset. This is an example of the narrative approach in which one uses qualitative sources to obtain information that helps identify particular shocks or other developments in the economy. Romer and Romer (1989) is a well-known example within modern macroeconomics.[1] Their study sought to identify exogenous changes in monetary policy by reading transcripts of Federal Reserve meetings in order to find dates at which policy was changed despite the fact that policymakers felt that economic activity was at an appropriate level.

An alternative natural experiment strategy exploits the fact that military spending does not occur uniformly across an entire country but is concentrated in some regions. Using records of military purchases to understand the locations of military suppliers, Nakamura and Steinsson (2014) are able to construct a measure of military spending at the level of U.S. states. They show an increase in national military spending leads to a disproportionately large increase in spending in some states as compared to other states. A fiscal multiplier can then be estimated from the output movements in states with more spending relative to the output change in the states with less spending. The identifying assumption is that there are no other shocks which are both correlated with military spending in the time series and correlated with the fiscal loadings in the cross section. Nakamura and Steinsson estimate a multiplier of 1.5. The regional data provide useful evidence on the effects of government spending on the economy. It is worth noting, however, that a multiplier across states is not directly comparable to an economy-wide multiplier. When government spending increases at the national level, there may be changes in national taxes, interest rates, or commodity

[1] See also Romer and Romer (2023) for a review and update of this work.

prices. But these general equilibrium effects are common across states and therefore do not contribute to the estimated regional multiplier. For example, if the war spending is financed by national taxes, the tax impact will be the same (or similar) across states and therefore will not be part of the regional multiplier but it will be part of the national multiplier. Nakamura and Steinsson (2014) argue that the large regional multiplier they find is consistent with Keynesian theories in which an increase in demand leads to an economic expansion because it is the increase in demand that differs across states.

Natural experiments are often compelling because they isolate changes in the economy that are plausibly exogenous. Oftentimes, this requires narrowing our focus to particular types of variation in the economy and we may worry that these events are not representative of the phenomenon as a whole. For example, it is possible that a general increase in government consumption has a different impact on the economy than military spending does. Moreover, there is no guarantee that we can find a convincing natural experiment for our question of interest. Therefore we may want to consider alternative empirical strategies.

8.4 Structural Vector Autoregressions

One option is to use the time series of aggregate government spending and ask how surprise innovations to this series affect the economy. The challenge here is that the surprise in spending may reflect an endogenous response of spending to some other development rather than a fiscal policy shock. The literature on structural vector autoregressions (VARs) seeks to overcome this challenge in order to uncover causal relationships in time series data.

Consider the process

$$B_0 Y_t = B_1 Y_{t-1} + B_2 Y_{t-2} + \cdots + B_J Y_{t-J} + \varepsilon_t, \tag{8.4}$$

where $Y_t \in \mathbb{R}^n$ is a vector of observed data, B_j is an $n \times n$ matrix of coefficients, and $\varepsilon_t \in \mathbb{R}^n$ is an i.i.d. stochastic innovation. Each equation in this system expresses a relationship between the current Y_t and lags of Y_t plus an error term. In a **structural vector autoregression** we give the equations and shocks a causal interpretation. We would like to think of one equation shifting without affecting other equations. For example, a TFP shock only shifts the production function without affecting other equations in the model.[2] We assume the structural shocks are uncorrelated with one another. It is without loss of generality to assume that the covariance matrix of ε_t is the identity matrix as we can rescale the B_j's by the standard deviations of the shocks.

To make things concrete, consider a log-linearized version of (8.2) and (8.3), assuming that the average value of η_t is 1,

$$\hat{y}_t = \frac{\chi}{1+\chi}\hat{G}_t + \frac{1}{1+\chi}\hat{A}_t \tag{8.5}$$

[2]In this context, the word "structural" is used in the manner it is used in the study of simultaneous equation models in econometrics. In the structural system, the equations and shocks are given a causal interpretation. In other contexts, the term "structural" has a different meaning and refers, instead, to parametric models that are explicitly derived from economic theory.

and

$$\hat{G}_t = \hat{y}_t + \frac{1}{1+\gamma}\hat{\eta}_t, \tag{8.6}$$

where

$$\chi \equiv \frac{1}{1+\psi}\frac{\bar{G}/\bar{y}}{1-\bar{G}/\bar{y}}.$$

The shock in each equation has a clear interpretation: $\hat{A}_t$ is a shock to the supply-side of the economy while $\hat{\eta}_t$ is a shock to fiscal policy. The equations in the system cannot be estimated directly due to endogeneity concerns. In the first equation, G_t is correlated with A_t because it is determined by the solution to the whole system, which depends on A_t. But if we can obtain consistent estimates of the parameters, we would have most of the information we need to calculate the fiscal multiplier because it tells us the causal effect of $\log G$ on $\log y$.[3]

While we cannot estimate the structural system (8.4) directly we can transform it by premultiplying by B_0^{-1} to obtain the **reduced-form** VAR

$$Y_t = \underbrace{B_0^{-1}B_1}_{A_1} Y_{t-1} + \underbrace{B_0^{-1}B_2}_{A_2} Y_{t-2} + \cdots + \underbrace{B_0^{-1}B_J}_{A_J} Y_{t-J} + \underbrace{B_0^{-1}\varepsilon_t}_{u_t}. \tag{8.7}$$

As we have now eliminated the contemporaneous variables from each equation, we can estimate the equations of (8.7) via OLS, but the shocks we recover are not the structural shocks ε. Instead, the reduced-form residuals u_t are linear combinations of the structural shocks with weights given by B_0^{-1}. If our goal is to understand how a shift in fiscal policy affects the economy, we need to isolate a fiscal policy shock so the reduced-form VAR itself does not allow us to answer the question of interest.

As we assumed $\text{Var}(\varepsilon_t) = I$, it follows that $\text{Var}(u_t) = B_0^{-1}B_0^{-1\prime}$. In this way, the estimated covariance matrix of u_t gives us information about B_0. However, it does not give us all the information we need. Let F be some orthogonal matrix, which means $F'F = I$. Then for any candidate solution b such that $bb' = \text{Var}(u_t)$ there is an alternative solution bF' that also satisfies $bF'Fb' = \text{Var}(u_t)$. To resolve this problem, we need more information that allows us to eliminate all these rotations F except for one. This information does not come from the estimated reduced-form VAR but from identifying assumptions.

The matrix $\text{Var}(u_t)$ has dimension $n \times n$ but it is symmetric so it only contains $n(n+1)/2$ distinct values. To determine all of B_0 we need an additional $n^2 - n(n+1)/2$ restrictions on B_0.[4] These restrictions are identifying assumptions that must be motivated by theory or other data besides that contained in Y_t.

In our example system (8.5)–(8.6), we have $n = 2$ so we need one restriction on B_0. Our theoretical model implies that the coefficient on $\hat{y}_t$ in (8.6) is equal to 1 and we could use that as our identifying assumption. How does restricting a coefficient to 1 solve the identification problem? By defining $\hat{g}_t \equiv \hat{G}_t - \hat{y}_t$, we could rewrite (8.5)–(8.6) as

$$\hat{y}_t = \chi\hat{g}_t + \hat{A}_t \tag{8.8}$$

[3] The system (8.5)–(8.6) is written in logs while the fiscal multiplier is defined in levels. We can convert from logs to levels by multiplying the estimated effect in logs by $\bar{y}/\bar{G}$. Appendix 8.A.2 verifies this in the context of the model here.

[4] We are simply counting equations and unknowns here: the unrestricted B_0 has n^2 unknowns but due to symmetry $B_0^{-1}B_0^{-1\prime} = \text{Var}(u_t)$ only provides $n(n+1)/2$ equations. We therefore need to impose additional equations ("restrictions") to determine B_0.

and

$$\hat{g}_t = \frac{1}{1+\gamma}\hat{\eta}_t. \tag{8.9}$$

While G_t is endogenous, equation (8.9) makes clear that our theory asserts that g_t is exogenous and we can then attribute any movements in g_t to fiscal shocks. There is then no endogeneity problem in equation (8.8) because g_t is exogenous.

Recursive identification

The system (8.8)–(8.9) is an example of **recursive** identification. There is one equation that has no contemporaneous variables on the right-hand side, which allows us to identify shocks to the variable on the left-hand side of that equation. We can then treat that variable as exogenous in other equations and identify shocks to equations that only have exogenous variables on the right-hand side. In a larger system we could potentially identify all the shocks in a series of steps.

A recursive identification scheme is also called "Cholesky" or "triangular" identification. Recall that a Cholesky decomposition of a real positive definite matrix M gives a lower-triangular matrix L such that $LL' = M$. If we order the variables in Y_t so that the first variable does not depend on any other contemporaneous variable, the second variable may depend on the first but no other contemporaneous variables and so on, then a Cholesky decomposition of the covariance matrix $\text{Var}(u_t)$ gives B_0^{-1}.

Equation (8.9) imposes the restriction that $\hat{g}_t$ does not depend on $\hat{y}_t$. In an application with lagged variables, we could allow $\hat{g}_t$ to depend on lags of $\hat{y}_t$. So the identifying assumption only restricts the immediate impact of the shocks and not the dynamic relationships between the variables. This is the appeal of structural VARs, we can allow the dynamics of the economy to be quite flexible.

When we impose that certain elements of B_0 are equal to zero, we call those "timing restrictions" because we are imposing that certain variables do not depend on others within the same period but they can depend on those other variables with a delay. Blanchard and Perotti (2002) use timing restrictions to identify fiscal policy shocks in a structural VAR. They assume that fiscal policymakers cannot immediately (within a quarter) observe current GDP and adjust fiscal policy because data on GDP only arrives after the quarter is complete and because it takes time to pass legislation to adjust policy. This argument motivates a specification in which the contemporaneous response of spending to GDP is restricted to zero. Blanchard and Perotti's results are somewhat sensitive to the details of the specification but the multipliers they find are generally around one.[5]

Timing restrictions are not the only route to identification in a structural VAR. One approach that is widely used now is to use a natural experiment as an instrumental variable to identify a shock of interest in the VAR. In another approach, Uhlig (2005) proposed imposing restrictions on the signs of the responses of some variables following the shock of interest. Alternatively, Blanchard and Quah (1989) used the identifying assumption that only

[5] Recall that Blanchard and Perotti define the multiplier as the ratio of the peak output response to the peak spending response.

supply shocks can affect GDP in the long-run while demand shocks can lead to temporary fluctuations—a so-called "long-run restriction."

From natural experiments to impulse response functions using VARs

Let z_t be an exogenous variable, say a measure of military spending news. One way of computing impulse responses to shocks to z_t is to include z_t as an exogenous variable in a VAR. Specifically, one could use a recursive identification scheme with the assumption that z_t does not respond to any of the other variables in the system (i.e. it is ordered first). In her study of military spending, Ramey (2011) estimates the impulse responses by including the military news as an exogenous series in a VAR. This procedure is an alternative to a local projection

A limitation of structural VAR modeling is the invertibility assumption that the reduced-form innovations at date t are linear combinations of the structural shocks at date t. This assumption is implausible because it asserts that the variables in the VAR and their lags summarize all of the relevant information about the state of the economy and the only reason the VAR has a forecast surprise at date t is because there is a structural shock at date t. In contrast, suppose the estimated VAR system omits some useful information about the state of the economy. As this information is ignored, there will be a forecast surprise at t that is not due to a structural shock at t, but rather a symptom of the imperfect information at $t-1$. In general, it is not possible to test invertibility—you cannot be sure there is no other useful information. Instead, one can attempt to include a rich set of variables that plausibly capture the main factors relevant to forecasting. On the other hand, the number of parameters to estimate grows rapidly as the system expands leading to overfitting concerns. Resolving this tension between a richer information set and overfitting has been, and continues to be, an important issue for research.

The appeal of structural VARs is that they impose relatively few assumptions on the dynamics of the economy. On the other hand, there are also benefits that come with making more structural assumptions in the form of a model built on microeconomic foundations. With a full description of the economic environment, we can be more specific about economic mechanisms, the welfare consequences of shocks and policy decisions, and we can connect data from a variety of sources including micro-level data.

8.5 VARs and local projections

We often want to compute an impulse response function (IRF) that summarizes how one variable responds dynamically to innovations in another or in itself. Two common ways of estimating IRFs are local projections and VARs. This section describes the relationship between these estimation techniques, the IRFs that they seek to estimate, and how the two methods compare.

Let y_t be an $n \times 1$ vector of variables at date t. Without loss of generality, we can suppose that we are interested in the response of y_t to a change in the variable in the first position of the vector, $y_{1,t}$. Before going further, it is worth briefly mentioning identification. Nothing we will discuss in this section has to do with a structural interpretation of innovations to $y_{1,t}$. If you have identified quasi-random variation in an instrument z_t you can simply include it as $y_{1,t}$ and use either method to estimate how outcomes of interest react to the instrument. Similarly, if you are interested in simply estimating dynamic relationships between variables without a structural interpretation you may again use either method.

A local projection constructs an IRF one horizon at a time. Letting $h \geq 0$ be the horizon, we consider the regression

$$y_{i,t+h} = \beta_{i,h} y_{1,t} + \sum_{j=1}^{J} A_{i,j} y_{t-j} + \varepsilon_{i,h}.$$

We run this regression for $h = 0, 1, \cdots, H$ and then the impulse response is $\{\beta_{i,h}\}_{h=0}^{H}$.

A (reduced-form) VAR specifies

$$y_t = A_1 y_{t-1} + A_2 y_{t-2} + \cdots + A_J y_{t-J} + u_t,$$

where the A's are $n \times n$ coefficient matrices and u_t is an $n \times 1$ vector of innovations. Each row of this system is a regression equation and we can estimate the A's one equation at a time using OLS. When $J = 1$, the IRFs are easy to calculate. The impulse responses at horizon h are given by A_1^h, where the (i, j) element gives the response of $y_{i,t+h}$ to $\varepsilon_{j,t}$. Notice that at horizon 0, $A_1^h = I$ so the reduced-form shocks only perturb a single variable on impact. This is also true when $J > 1$, but now the dynamic IRFs are more complicated. In those cases, we can construct the companion form that expresses the system as a first-order system in an expanded set of variables. For example for $J = 2$ the companion form is

$$\mathcal{A} = \begin{pmatrix} A_1 & A_2 \\ I & 0 \end{pmatrix}$$

with

$$\begin{pmatrix} y_t \\ y_{t-1} \end{pmatrix} = \mathcal{A} \begin{pmatrix} y_{t-1} \\ y_{t-2} \end{pmatrix} + \begin{pmatrix} u_t \\ 0 \end{pmatrix}.$$

The upper-left $n \times n$ block of $\mathcal{A}^h$ then gives the IRFs at horizon h. As before, the (i, j) element gives the response of $y_{i,t+h}$ to $u_{j,t}$.

These two estimation techniques as described above are estimating different things. If we assume $\{y_t\}$ is Gaussian, so least squares equates with a conditional expectation, the local projection will estimate

$$\mathbb{E}\left[y_{i,t+h} | y_{1,t} = 1, \{y_{t-j}\}_{j=1}^{J}\right] - \mathbb{E}\left[y_{i,t+h} | y_{1,t} = 0, \{y_{t-j}\}_{j=1}^{J}\right].$$

Notice that these expectations do not condition on $y_{j,t}$ for $j \geq 2$. In contrast, the (reduced-form) VAR assumes that there is no contemporaneous impact of one variable on another and is therefore estimating

$$\mathbb{E}\left[y_{i,t+h} | y_t = \begin{pmatrix} 1 & 0 & \cdots & 0 \end{pmatrix}', \{y_{t-j}\}_{j=1}^{J}\right] - \mathbb{E}\left[y_{i,t+h} | y_t = \begin{pmatrix} 0 & 0 & \cdots & 0 \end{pmatrix}', \{y_{t-j}\}_{j=1}^{J}\right].$$

Plagborg-Møller and Wolf (2021) showed the IRFs from the reduced-form VAR can be combined to give the same impulse responses as the local projection.[6]

The local projection is very flexible in how it estimates impulse responses whereas the VAR imposes the assumptions of the statistical model to extrapolate the longer-horizon response from the short-term dynamics. With many lags, the VAR can be as flexible as the local projection, but in typical practice one has a moderate number of lags in a VAR.

The local projection is regressing y_{t+h} on $y_{1,t}$ (and lags). As h grows, there are many other shocks that have occurred in the meantime so the signal to noise ratio can be small leading to wide standard errors. The VAR, in contrast, regresses y_{t+1} on y_t (and lags). The VAR impulse responses are then extrapolated from these one-step-ahead dynamics. As the horizon h grows, the uncertainty around our estimate usually does grow at least for moderate h but less than for the local projection. So, the VAR will typically yield estimates with tighter confidence bands. However, if the auto-covariance function of the data is not well approximated by a VAR(J), the mis-specification of the statistical model can lead to larger bias in the VAR estimates than in the local projection estimates. In total, there can be a bias-variance tradeoff when we choose between a VAR and a local projection.

Figure 8.3 illustrates the bias-variance tradeoff. To generate the figure, we simulate a VAR process with two lags and then estimate local projections and a VAR with only one lag. We repeat the simulation 1,000 times and plot the 5th percentile, median, and 95th percentile of our estimated impulse responses. The dashed lines in the top panel show the true impulse responses from the VAR(2). In the bottom panel, we subtract the true IRF for ease of comparison. Clearly the local projection has more sampling variation but the VAR has more bias. For this example, the mean square errors are 0.0047 and 0.0027 for the local projection and VAR, respectively.

How to choose between the methods depends on your priorities. For point estimates, Li, Plagborg-Møller, and Wolf (2024) conducted a simulation study and found that for a researcher that wishes to estimate impulse responses with low mean square errors, the VAR approach is more appealing for many data generating processes. However, for inference, the bias in VAR estimates implies that standard confidence intervals typically do not cover the true impulse responses with the probability they are meant to (see Olea, Plagborg-Møller, Qian, and Wolf, 2024).

8.6 A model of fiscal policy

The final two empirical strategies we will discuss make use a of a structural model. A model can be viewed as an analytical tool that allows us to make sense of data and reach conclusions that are not immediately clear from the data alone. This is both an appeal of using models and a concern with using models. In some cases, the data do not speak

[6] The logic here is that we need to find the expected innovations to $y_{j,t}$ for $j \geq 2$ given a unit innovation to $y_{1,t}$. We then use these expected innovations in the other series as weights to combine the IRFs for $j = 1, \cdots, J$. Let Σ be the covariance matrix of the reduced-form VAR residuals. Let L be the lower-triangular Cholesky decomposition of Σ such that $LL' = \Sigma$. Now, consider the first column of L, call it $L_{\cdot,1}$ and normalize so that the first element of $L_{\cdot,1} = 1$. This vector gives the weights for a linear combination of the reduced-form VAR innovations that replicates the local projection estimand.

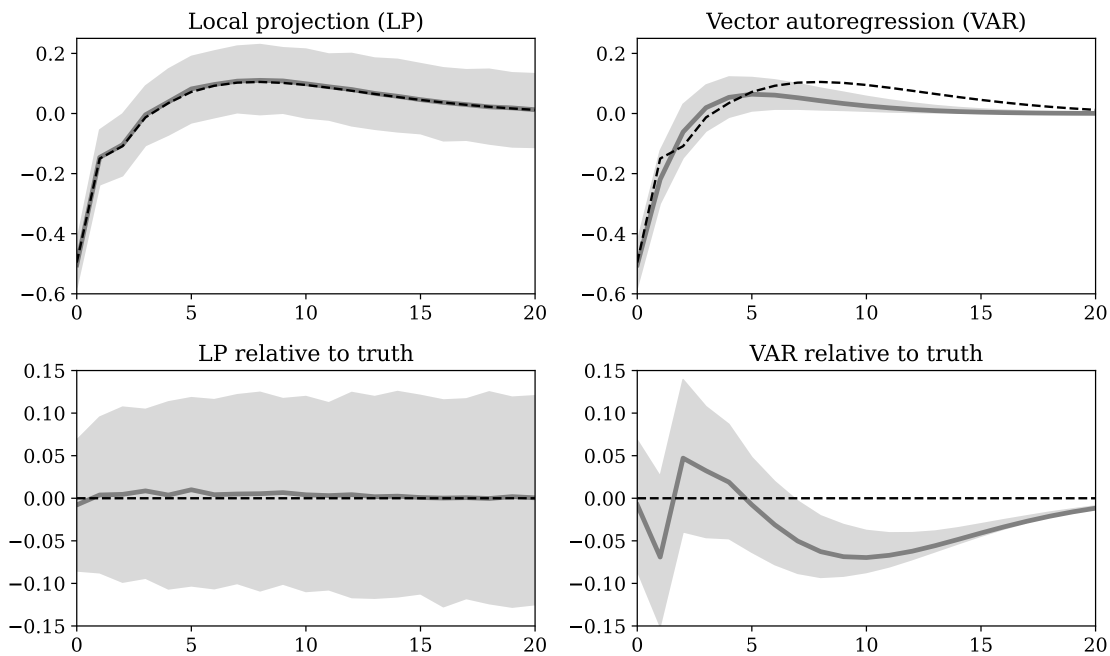

Figure 8.3: Bias-variance tradeoff in simulated data

Notes: We simulated 1,000 draws of length $T = 200$ from a VAR(2) in two variables. For each simulation draw, we estimate the impulse responses using local projection and a VAR but with only one lag. The solid lines are the median impulse responses at each horizon across the 1,000 draws. The shaded areas show the 5th and 95th percentiles at each horizon. The dashed lines show the true impulse responses. The lower panels show the differences between the estimated impulse responses and the true values.

directly to the question of interest because the type of experiment we are interested in has never occurred. We then have no choice but to use assumptions, typically in the form of a model, to extrapolate from the data we have observed to the question we are interested in. On the other hand, we may worry that the model assumptions drive our conclusions without a solid grounding in facts. To alleviate this concern, we need to use a model that allows for the economic channels that are most relevant to our question and allows the data to determine the relative strengths of these channels.

In our analysis of the fiscal multiplier, the static model presented in Section 8.2 only allows for one choice for the economy after an increase in government spending: the economy can work more or consume less. If we allow for investment, there is another possibility: the economy could reduce investment to free up resources for the government. In this section we will expand our model to incorporate this margin of adjustment. While the expanded model is a step in the right direction, it omits the distortionary effects of taxes and it omits Keynesian channels that are sometimes considered important in determining the effects of fiscal policy.[7]

8.6.1 The model

The main change we will now make relative to Section 8.2 is to include capital and investment. We will also now analyze a competitive equilibrium rather than a planner's solution because the model's implications for prices and payments to capital and labor are useful when calibrating the model.

A representative household has the preferences

$$\mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t U_t,$$

where U_t is given in (8.1). Output is produced according to the production function

$$y_t = A_t k_t^{\alpha} \ell_t^{1-\alpha}. \tag{8.10}$$

Goods can be used for consumption by households, consumption by the government, or investment. Letting I_t denote investment, capital accumulates according to $k_{t+1} = (1-\delta)k_t + I_t$ and the aggregate resource constraint is $y_t = c_t + I_t + G_t$. Combining the two we have

$$y_t = c_t + k_{t+1} - (1-\delta)k_t + G_t. \tag{8.11}$$

For a given level of production, an increase in G_t will reduce the resources available for private consumption and investment.

We assume that households accumulate capital and rent it and their labor to firms in spot markets at rental rate r_t for capital and wage w_t for labor. Households must pay lump-sum taxes in amount T_t. The representative household's budget constraint is

$$c_t + k_{t+1} = (1 + r_t - \delta)K_t + w_t \ell_t - T_t.$$

[7] Chapter 15 discusses tax distortions and Chapter 18 will discusses new Keynesian models of the business cycle.

The government uses the tax revenue to finance its purchases so the government budget constraint is $T_t = G_t$. Imposing taxes on the household will affect their choices as they must consume less, save less, or work more in order to pay the taxes. However, as the taxes are imposed lump-sum, they do not affect the marginal returns to work and saving and therefore do not distort the household's choices away from the socially efficient choices.

There are two exogenous processes in this economy, which we assume follow AR(1) processes with unit mean:

$$A_t = (1 - \rho_A) + \rho_A A_{t-1} + \varepsilon_{A,t} \tag{8.12}$$

and

$$\eta_t = (1 - \rho_\eta) + \rho_\eta \eta_{t-1} + \varepsilon_{\eta,t}, \tag{8.13}$$

where $\varepsilon_{A,t}$ and $\varepsilon_{\eta,t}$ are i.i.d. with zero mean and variances σ_A^2 and σ_η^2, respectively.

The household's problem is to choose consumption, savings and labor supply to maximize its utility subject to its budget constraint taking the prices and taxes as given. The solution to this problem follows the same steps as in Section 7.5 so we will simply state the optimality conditions:

$$\frac{1}{c_t} = \beta \mathbb{E}_t \left[(1 + r_{t+1} - \delta) \frac{1}{c_{t+1}} \right] \tag{8.14}$$

and

$$\frac{w_t}{c_t} = \ell_t^\psi. \tag{8.15}$$

Equation (8.14) is the consumption Euler equation under uncertainty. Equation (8.15) is the first-order condition for labor supply that equates the utility of working an extra unit of time and consuming the income to the marginal disutility of working more.

The representative firm's problem is to maximize profits by hiring labor and capital and using these inputs to produce. As usual, the firm's first-order conditions equate marginal products and input prices:

$$r_t = \alpha \frac{y_t}{k_t} \tag{8.16}$$

and

$$w_t = (1 - \alpha) \frac{y_t}{\ell_t}. \tag{8.17}$$

We will assume that the government chooses taxes and spending to maximize the welfare of the representative household. In doing so, the government can anticipate that taxes cause households to change their behavior. Chapter 15 will discuss this type of interaction between the government and the private sector in more detail. Here we note that the solution to the government's problem satisfies the same first-order condition as the planner's problem, which equates the marginal benefit of spending to the marginal cost of the household consuming less

$$\frac{\gamma \eta_t}{G_t} = \frac{1}{c_t}. \tag{8.18}$$

An equilibrium of this economy is a set of stochastic processes $\{k_t, c_t, y_t, \ell_t, G_t, r_t, w_t, A_t, \eta_t\}$ that satisfy equations (8.10)-(8.18). In what follows we will discuss functional forms and parameter values for the model. Once those elements are in place, we can compute a solution to the model via several methods. To analyze the dynamics of the economy, we will use a linear approximation around the deterministic steady state as discussed in Section 7.3.4.[8]

8.6.2 The fiscal multiplier

To develop intuition for the fiscal multiplier in this model, we can begin by supposing that investment does not respond to the change in spending. As investment is unchanged, there are only two ways to free up the resources for the extra spending: work more or consume less. If the economy simply consumes less, then output is unchanged and the fiscal multiplier will be zero. On the other hand, if the economy simply works more, then output increases one-for-one with spending and the multiplier will be one. From the perspective of households, an increase in government spending means an increase in taxes and therefore it induces a negative wealth effect that causes them to consume less and work more so the actual outcome is somewhere between these two extremes and the fiscal multiplier will be between zero and one.

If we allow investment to respond, there is a third way to pay for the government spending: invest less. If the economy invests less, it can free up resources for the government but at the cost of having a smaller capital stock in the future. With less capital in the future, output is lower. By reducing investment the households have extra resources that they can use to consume more and work less. To the extent they work less, the fiscal multiplier will be lower. If the change in government spending is short-lived, reducing investment will help to smooth consumption—the economy reduces the capital stock to finance the government and then builds the capital stock back up once the government spending returns to normal. However, if the spending shock is highly persistent, then this "borrow from the future" strategy is not as effective because the economy needs to finance the government now and in the future. Therefore the fiscal multiplier is increasing in the persistence of the shock to spending.

8.7 Structural estimation

Structural estimation treats the economic model as a statistical model of the data we observe. For different configurations of the structural parameters, the observed data are more or less likely to have occurred. Using likelihood methods (maximum likelihood or Bayesian estimation) we can then estimate the parameters of the model from the observed data. To start, we will continue with the simplified model from Section 8.2 before generalizing to the richer model from Section 8.6.

To estimate the model with likelihood methods we need to make assumptions about how the shocks are distributed. For the sake of discussion, let us suppose A_t and η_t are i.i.d. with densities p_A and p_η. Suppose further that our observable variables are time series of government spending and output from dates 1 to T. We then ask how likely these data were for a given set of parameters. For a given date t, we can use (8.2) and (8.3) to solve for the

[8]See Appendix 8.A.1 for how to calculate the steady state of this model.

values of A_t and η_t that are implied by the observed y_t and G_t

$$A_t = y_t \left(1 - \frac{G_t}{y_t}\right)^{\frac{1}{1+\psi}} \tag{8.19}$$

and

$$\eta_t = \frac{G_t/y_t}{\gamma(1 - G_t/y_t)}. \tag{8.20}$$

The likelihood of these observations is then $\text{Prob}(G_t, Y_t|\psi, \gamma) = p_A(A_t) \times p_\eta(\eta_t)$ and the likelihood of the full sample is then

$$\text{Prob}(\{G_t, Y_t\}_{t=1}^T |\psi, \gamma) = \prod_{t=1}^T (p_A(A_t) \times p_\eta(\eta_t)).$$

Maximum likelihood estimation treats the likelihood function above as the objective function to maximize. Bayesian estimation uses Bayes rule to derive a posterior probability distribution of the parameters conditional on the data. Doing so requires a prior distribution of our beliefs over the parameters before observing the data. Once we have estimated the parameters, we can differentiate (8.2) to find the implied fiscal multiplier (see Appendix 8.A.2 for a derivation of the fiscal multiplier in this model).

It is worth considering how structural estimation resolves the identification challenge discussed above. Equation (8.3) imposes a functional form on the fiscal rule that implies the G/y ratio is not affected by TFP shocks. Therefore, any changes in G/y are due to fiscal shocks. We are then able to infer the fiscal shocks from movements in G/y as shown in equation (8.20). A reasonable criticism of this approach is that the functional form assumption for the fiscal rule may be mis-specified, leading to mis-identification of the fiscal shocks.

Calculating the likelihood for the static model is straightforward because we can solve for η_t and A_t directly from the model equations. In actual applications this typically won't be possible in part because there are unobserved state variables. For example, with the production function $y_t = A_t k_t^\alpha \ell_t^{1-\alpha}$, if k_t and A_t are unobserved, then we cannot say if output per hour is high because A_t is high or because k_t is high. The process of using observed data to infer the position of an unobserved state is known as filtering.

We will give a brief overview of how one can use filtering techniques to construct the likelihood for the expanded model with capital. The model has three state variables (k_t, A_t, and η_t). We do not observe A_t or η_t. While we have measures of the capital stock, they are likely contaminated by substantial measurement error so we often treat k_t as unobserved, too. We will suppose we can observe the time series of y_t and G_t.

The solution of the model provides a policy rule for investment in the form

$$k_{t+1} = f(k_t, A_t, \eta_t).$$

This rule, along with the laws of motion for the stochastic processes (8.12) & (8.13), define a stochastic difference equation for the state vector $X_t \equiv (k_t, A_t, \eta_t)'$

$$X_{t+1} = F(X_t, \varepsilon_{t+1})$$

where $\varepsilon_{t+1} \equiv (\varepsilon_{A,t+1}, \varepsilon_{\eta,t+1})'$ is a vector of shocks. Similarly, we can use the other equations of the model to express the observed data $Y_t \equiv (y_t, G_t)'$ in terms of the states and we can use $Y_t = H(X_t)$ to represent this relationship. The functions F and H are non-linear functions that depend on the model parameters. If we linearize these equations we obtain

$$\hat{X}_{t+1} = \mathcal{A}(\theta)\hat{X}_t + \mathcal{B}(\theta)\varepsilon_{t+1} \tag{8.21}$$

$$\hat{Y}_t = \mathcal{C}(\theta)\hat{X}_t, \tag{8.22}$$

where $\mathcal{A}(\theta)$, $\mathcal{B}(\theta)$, and $\mathcal{C}(\theta)$ are coefficient matrices that depend on the model parameters, which we collect in the vector θ, and $\hat{X}$ is the deviation of X from steady state.

It is common to assume that ε is a Gaussian random variable because this allows us to use a particularly convenient filtering algorithm: the Kalman filter. We will specifically assume $\varepsilon \sim N(0, I)$. The Kalman filter generates estimates of the state at date t give information through $t-1$. These estimates are uncertain and the linear-Gaussian framework implies $X_t \sim N(X_{t|t-1}, P_{t|t-1})$, where $X_{t|t-1}$ is the conditional mean of X_t given information through date $t-1$ and $P_{t|t-1}$ is the covariance matrix that reflects our uncertainty over X_t. Given (8.22), we then have a distribution over Y_t that is also Gaussian with a mean $\mathcal{C}(\theta)X_{t|t-1}$ and covariance matrix $\mathcal{C}(\theta)P_{t|t-1}\mathcal{C}(\theta)'$. This gives us the likelihood of Y_t given data through $t-1$. We then can form the likelihood of full sample by applying this same logic to all dates t:

$$\text{Prob}(\{Y_t\}_{t=1}^{T}|\theta) = \prod_{t=1}^{T} \mathcal{N}\left(Y_t|\mathcal{C}(\theta)X_{t|t-1}, \mathcal{C}(\theta)P_{t|t-1}\mathcal{C}(\theta)'\right),$$

where $\mathcal{N}(Y|\mu, \Sigma)$ is the density of multivariate normal distribution with mean μ and covariance Σ evaluated at Y.[9] Once we have constructed the likelihood we can proceed with maximum likelihood or Bayesian estimation of the parameter vector θ.

To review, we have described a widely-used method for structurally estimating macro models. For a given θ, one solves for a linear approximation to the dynamics of the economy. One then uses the Kalman filter to filter the observed data, estimate the unobserved state, and construct the likelihood. Use of the Kalman filter assumes the shocks to the model are Gaussian. There exist methods for handling non-linear and non-Gaussian models, but the tractability of the Kalman filter leads linear-Gaussian models to be used in many applications.

Structural estimation treats the model as the data generating process. We next turn to calibration, where the goal is not to fully explain the data, but rather the gauge the strength of the key mechanisms embedded in the model.

Limited-information estimation

In this section, we have focused on likelihood methods or "full-information" estimation because that is a very common approach to structural estimation within macroeconomics. However, one could also consider limited-information methods that use par-

[9] Notice that we require a belief about the initial distribution of X_1 given no data (i.e. $X_{1|0}$ and $P_{1|0}$). In stationary models, it is common to assume that the initial state is drawn from the system's invariant distribution.

ticular moments of the data as estimation targets. One might estimate a structural parameter by using a single equation from a larger model and using the generalized method of moments. Such an approach still requires identification to obtain consistent estimates of the parameters and a benefit of specifying the full model is that the argument for identification can be laid out in concrete terms. Alternatively, estimated impulse response functions can be used as estimation targets. This procedure, known as impulse response matching, can be a bridge between the structural VAR or natural experiment approaches and fully structural modeling.

8.8 Calibration

Suppose we have some data of interest. This could be simple moments we calculate such as the volatility of GDP or it could be the result of some other analysis such as a study of a natural experiment. We can then ask if the mechanisms embedded in the model are able to account for this data of interest. To answer that question, we need parameters for the model. One option is to use the data of interest in picking the parameters, but this is not so compelling as it only tells us that there exist some parameters that can explain the data. Instead, we might seek other sources of information about the parameters. Then, with the parameters already set, we compare the model to the data of interest. This is the approach taken in calibration. In calibrating a model, we must be clear about what mechanisms our model uses to explain the data of interest and then take care to choose the parameters of the model with information that speaks to the strength of the relevant mechanisms.

In our analysis of fiscal policy, the data of interest could be the empirical estimates discussed above, which put the fiscal multiplier near one. Section 8.6 highlighted several considerations that are important to the size of the fiscal multiplier. The multiplier will be larger if households are more willing to vary their labor supply than their consumption. The fiscal multiplier will also be larger if the economy is unwilling or unable to reduce investment. As we calibrate the model, we should pay special attention to the parameters that affect these channels as they ultimately will determine the fiscal multiplier.

Chapter 3 has already discussed a calibration strategy for some of the parameters in our model. The parameter α determines the share of aggregate income that accrues to capital, which we observed to be in vicinity of 1/3. Chapter 3 also uses the average ratio of investment to capital, which was found to be 0.076, to calibrate the depreciation rate. The model presented here does not have population growth or trend productivity growth, so the only reason to invest in steady state is to replace the depreciated capital. The 0.076 ratio is for annual investment (a flow variable e.g. the investment spending over the course of a year) relative to the value of the capital stock (a stock variable e.g. the value at the start of the year). For the analysis of short-run dynamics, we often specify a model time period to be one quarter of a year. In quarterly terms, the investment-capital ratio is $0.076/4 = 0.019$ and we will set the depreciation rate $\delta = 0.019$.[10] From the Euler equation, we can

[10]The true depreciation rate is lower as some of the observed investment is explained by a growing economy. One could then argue that we should set δ lower, but we chose not to do that. If the economy invests less to finance government spending, it will move away from its steady state capital-output ratio. In reality, the decline in the capital-output ratio will reflect both depreciation and a growing economy. We will capture

see that the steady state return to capital is determined by $1/\beta$. Using (8.14) and (8.16) we obtain $\beta = (1 + \alpha\bar{y}/\bar{k} - \delta)^{-1}$. Inserting the observed capital-output ratio[11] of 3.3×4 into this equation yields $\beta = 0.994$.

The preceding parameters are quite closely connected to empirical counterparts, which allows for a straightforward calibration strategy. The inverse elasticity of labor supply, ψ, is more difficult to calibrate. The Frisch elasticity of labor supply is the percentage change in labor supply following a percentage change in the wage holding the marginal utility of consumption constant. With the preference specification used in the model, the Frisch elasticity is exactly $1/\psi$. A natural way to calibrate this parameter is to look for data on transitory changes in wages and see how much labor supply reacts. Wage fluctuations can reflect changes in labor supply as well as changes in labor demand. To identify the slope of the supply curve, we need to isolate changes in wages that are not themselves a reflection of labor supply shocks. One line of literature uses changes in tax rates as a source of variation in after-tax wages. This literature tends to find relatively small labor supply elasticities on the order of $1/2$ so $\psi \approx 2$ (see Chetty, 2012).[12]

In our model, a key mechanism is that the government imposes a lump-sum tax on households that makes them poorer and may cause them to work more through a wealth effect. The parameter ψ also determines the strength of this wealth effect in addition to the Frisch elasticity. We therefore want to be sure that our choice of ψ implies a reasonable magnitude of the implied wealth effect on labor supply. The ideal evidence on wealth effects on labor supply would involve random variation in wealth unrelated to other developments in the economy. Changes in wealth due to aggregate events, like an economic boom, are not ideal because the change in the economy that led to the boom may have also affected other aspects of the economy such as wages. Several studies have looked at the labor supply of individuals who win lottery prizes. Golosov, Graber, Mogstad, and Novgorodsky (2021) use data on lottery winners and compare their labor supply before and after they win. The authors find that winning \$100 leads, on average, to a \$2.30 drop in annual earnings. What does this imply for ψ? Solving equation (8.15) for ℓ and differentiating with respect to c we obtain $\frac{d\ell}{dc} = -\frac{1}{\psi}\frac{\ell}{c}$. We can multiply both sides by $4w$ to obtain the change in annual earnings. If we assume households smooth their consumption perfectly, they will consume the annuity value of their extra wealth. So receiving \$100 causes them to increase their consumption by $r/(1+r) \times \$100$. Putting these pieces together we have

$$\text{change in annual earnings} = -\$2.30 = -\frac{4}{\psi}\frac{w\ell}{c}\frac{r}{1+r}\$100.$$

We will use a quarterly interest rate of $r = 1\%$. The ratio $w\ell/c$ is quarterly earnings over quarterly consumption. Among households that work, this ratio is somewhat above 1 due to the life-cycle savings profile and we will set it to 1.2. Solving for ψ, we find $\psi = 2.1$. This is reassuring in that a value of ψ near 2 is consistent with evidence on the Frisch elasticity and with evidence on wealth effects.

---

both forces here by having a higher depreciation rate.

[11] As discussed in Chapter 2, the value of the capital stock relative to annual GDP is about 3.3. We multiply by 4 to express the ratio in terms of quarterly output rather than annual output.

[12] Some take the view that aggregating across a population of households changes the effective labor supply elasticity for the representative household. We will discuss this possibility in Chapter 14.

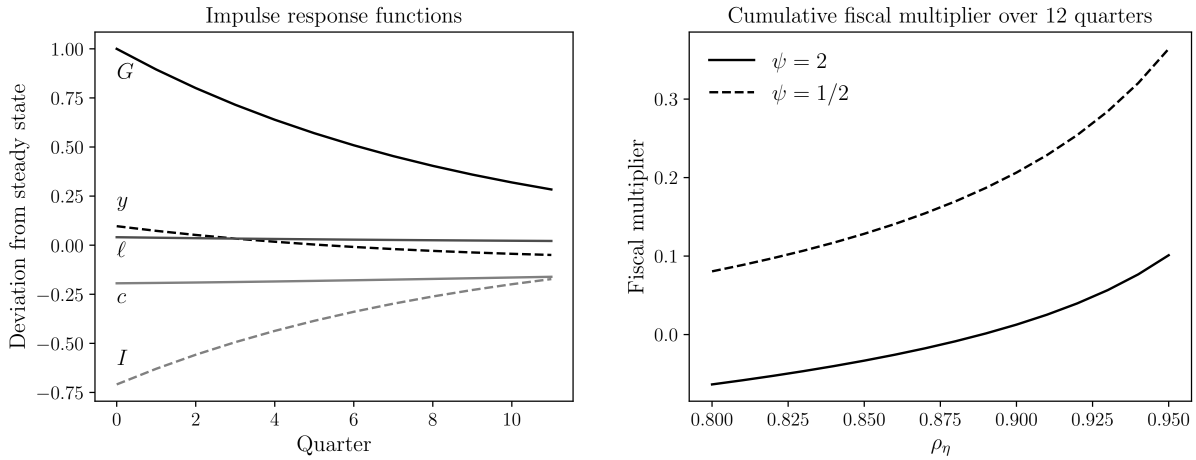

Figure 8.4: Impulse responses following a fiscal shock (left). Cumulative fiscal multiplier over the first 12 quarters (right).

In steady state, (8.18) becomes $\gamma = \bar{G}/\bar{c}$. In U.S. data from 1994 to 2023, the ratio of public consumption to private consumption has fluctuated in the range of 25% to 30%, while it was higher in the past (see Chapter 15 for data on the government's role in the economy). We will set $\gamma = 0.3$.

The model specifies exogenous stochastic processes for A_t and η_t. As our goal is to quantify the fiscal multiplier, we are interested in the economy's response to an exogenous change in fiscal policy, which we will generate with a shock to η_t. The presence of technology shocks will lead to consumption volatility and generate precautionary savings motives, which will have a small bearing on the fiscal multiplier. When we use a linear approximation of the model equations to analyze the economy, we implicitly assume these precautionary motives away. As a result, the economy's response to a fiscal shock will be the same regardless of whether or not technology shocks are included in the model. Additionally, the linear solution scales one-for-one with the size the shocks so the magnitude of the fiscal shock will not affect the fiscal multiplier. It follows that we do not need to specify ρ_A, σ_A^2 or σ_η^2. The fiscal multiplier does, however, depend on the persistence of the fiscal shock ρ_η. Our approach will be to report the fiscal multiplier for a range of values of ρ_η.

The left panel of Figure 8.4 shows the impulse responses to a fiscal shock (a positive realization of ε_η).[13] We have assumed that the shock is fairly persistent ($\rho_\eta = 0.9$) and G_t remains elevated for several years. In response to the shock, the economy works more, consumes less, and invests less. Quantitatively, we see that the bulk of the adjustment comes through investment and the change in labor supply is quite small. Initially, the increase in labor supply raises output. Over time, the decline in investment reduces the capital stock and, after 5 quarters, output falls below steady state. The cumulative fiscal multiplier—integrating the output and spending responses over the first 12 quarters—is close to zero.

In a sensitivity or robustness analysis we consider alternative calibrations and document how our conclusions are or are not affected by reasonable changes in the parameters. The right panel of Figure 8.4 shows that this conclusion is sensitive to the persistence of the

[13]We have scaled the shock so that G_t rises by one unit on impact, but this is just a normalization given the linear solution method.

spending shock. As ρ_η increases, reducing investment is less effective in smoothing consumption leading labor supply to adjust more resulting in a larger multiplier. Similarly, if ψ is smaller, so labor supply is more elastic, more of the adjustment occurs through labor supply and the fiscal multiplier is larger.

Model validation

Suppose that you have calibrated your model, but you still have additional empirical evidence that is relevant to your question. How can you incorporate this extra evidence into your analysis? One option is to ask whether your calibrated model can also match the additional facts. This step is called "model validation." The evidence employed in model validation is often related to the question of interest and perhaps less closely related to specific parameters. The point of model validation is to determine whether the model's answers can be trusted.

Earlier in the chapter we found fiscal multipliers closer to one while here we find multipliers closer to zero. From this analysis, it appears that wealth effects are unable to explain fiscal multipliers as large as those in the empirical estimates. We should note that this conclusion depends on the full model and it may be that wealth effects could be more powerful within a different model. Another possibility is that the model we consider here omits certain channels that make fiscal policy more powerful in stimulating output. The Keynesian view of fiscal policy is that it stimulates household incomes leading consumption to increase. This view contrasts quite strongly with the neoclassical view where government spending induces a negative wealth effect and a reduction in consumption.[14]

To conclude this section, we can draw some general principles for calibration. (i) *Understand the economics*: In calibrating a model, one has to make choices and the guiding principle is to obtain an accurate account of the strength of the main forces at play in the analysis. In order to make these choices, one must understand the economics of the problem. (ii) *Discipline the mechanisms*: Wherever possible, one should seek evidence that speaks directly to the strength of the mechanisms at the heart of the analysis. Microeconomic studies that cleanly isolate quasi-experimental variation in the economy are often an attractive source of information about the strength of mechanisms. (iii) *Robustness*: It is important to determine, and report, to what extent the quantitative results are sensitive to the parameter values chosen and to the exact way in which the model formalizes the mechanism in focus.

The conclusions one draws from a calibrated model are of course shaped in large part by the model itself. This is by design. Thus, the calibration example here asks about how large a fiscal multiplier would be if one only considered the mechanisms captured by the model at hand. The question is thus a partial one, to the extent one thinks the model lacks important elements. For example, if one thinks a Keynesian demand mechanism could be important then a next step is to add elements to the model, again by careful selection of parameters

[14] Chapter 18 discusses New Keynesian theories of the business cycle, which allow an increase in the demand for goods to increase output. Galí, López-Salido, and Vallés (2007) develop a New Keynesian model of fiscal policy and obtain fiscal multipliers near one.

relevant to the demand mechanism and evaluate again.

8.9 Taking stock

We have described some of the main methods that macroeconomists use for empirical and quantitative analysis. A reasonable question is "which one should I use?"

Macroeconomists increasingly value natural experiments as sources of identification. These natural experiments could occur at the national level, as in the military spending example above, or at a micro-economic level, as in the lottery winner example discussed in Section 8.8. In many cases, the natural experiment does not directly answer the question of interest, but it provides useful information about the economy that we can use as a calibration target for a model.

The strength of VAR analyses is that they place fewer restrictions on the dynamics of the economy than do full structural models. For example, in Figure 8.4 we see that in our structural model a government spending shock induces an immediate increase in output that then fades and a persistent decrease in consumption. These patterns are inherent to the model. If we calibrate or structurally estimate the model, the data will not be able to change these conclusions. In contrast, a VAR would allow more flexibility for the data to determine the shapes of the impulse responses.

Models take center stage when we lack data because they allow us to answer questions for which we do not have direct evidence. A model can serve as a bridge between the data we observe and the questions we want to answer. Our discussion of calibration is a good example—we can use data on capital, investment, and output from the national accounts along with data on how labor supply responds to taxes and lottery winnings to investigate a question about government spending. Models are also central when we are interested in understanding mechanisms at work in the economy.

It is worth noting that researchers often use multiple empirical methods in the same study. Natural experiments are often used to establish empirical results that may be explained by a structural model or may serve as a calibration target for a structural model. Regardless of which method you use, the success of your argument hinges on making a clear and compelling connection between your conclusion and the evidence that supports it.