6 The Relationship between VAR Models and Other Macroeconometric Models

This chapter puts the development of the VAR model in historical context and clarifies its relationship with other modeling frameworks used in empirical macroeconomics including dynamic simultaneous equations models (DSEMs) and dynamic stochastic general equilibrium (DSGE) models.

6.1 The Relationship between VAR Models and Traditional Dynamic Simultaneous Equations Models

Vector autoregressions as a tool for macroeconomic analysis were first introduced by Sims (1980a; 1980b). They quickly became the workhorse model in empirical macroeconomics as well as in empirical finance. Prior to 1980, the dominant model used in empirical macroeconomics was the DSEM. This class of models was developed by the Cowles Commission for Research in Economics (later renamed the Cowles Foundation), which concerned itself with linking economic theory to mathematics and statistics. Starting in the 1950s, macroeconomists at the Cowles Commission, including the subsequent Nobel Prize laureate Lawrence Klein, began constructing the first DSEMs of the U.S. macro economy (see Klein 1950; Klein and Goldberger 1955). Their work built on the pioneering work of Tinbergen (1939) and was inspired by the theoretical insights of Keynes' work. The objective of these researchers was to give quantitative content to Keynes' macroeconomic theory and to turn the DSEM into a tool for policymakers. Over time, DSEMs grew larger as economists strived for greater realism and sought to make the model more relevant for policy questions. These DSEMs often contained dozens or even hundreds of equations. They were driven by a large number of exogenous model variables, while including only a small number of endogenous model variables.

Exogeneity here is defined as the absence of contemporaneous or lagged feedback from the endogenous model variables to the exogenous model variables. Suppose that a vector of observed variables y_t may be partitioned into two classes of variables, 'exogenous', x_t, and 'endogenous', z_t. Then, following Zellner and Palm (1974), a structural dynamic, multivariate simultaneous linear system may be written as

$$\begin{bmatrix} H_{11}(L) & H_{12}(L) \\ 0 & H_{22}(L) \end{bmatrix} \begin{pmatrix} z_t \\ x_t \end{pmatrix} = \begin{bmatrix} F_{11}(L) & 0 \\ 0 & F_{22}(L) \end{bmatrix} \begin{pmatrix} w_{1t} \\ w_{2t} \end{pmatrix},$$

where $H_{ij}(L)$ and $F_{ij}(L)$, $i, j \in \{1, 2\}$, are matrix polynomials in the lag operator L and the elements of $w_t = (w'_{1t}, w'_{2t})'$ are mutually uncorrelated white noise. Such systems are known as structural vector ARMAX (or VARMAX) models, where the X refers to the set of strictly exogenous variables x_t. For further discussion of the concept of strict exogeneity, see Chapter 7. Conditions for the identification of the model parameters are discussed in Zellner and Palm (1974) and Hannan and Deistler (1988), for example. In practice, the structural equations are typically estimated by single-equation methods.

The implied reduced form of this system is obtained by expressing z_t as a function of lagged endogenous and current and lagged exogenous variables. Zellner and Palm (1974) show that this reduced-form representation takes the form of a VARMA (or, alternatively, VARMAX) model. The final form of this model is obtained by inverting the autoregressive lags for z_t and expressing z_t as a function of current and lagged values of the shocks and the exogenous variables only (see Zellner and Palm 1974; Wallis 1977; Lütkepohl 2005).

One advantage of large-scale DSEMs is that they allow detailed analysis of alternative policies and scenarios involving changes in the path of the exogenous model variables. Another advantage is that their economic structure facilitates the communication of the results to policymakers. For that reason, descendants of these models survive to this day at some private forecasting firms. Their demise, at least in academic research, came when they failed to explain the poor macroeconomic performance of the 1970s. This failure seemed to support the Lucas critique of structural DSEMs in the 1970s, which questioned the invariance of the estimated DSEMs to policy interventions (see Lucas 1976). It was followed by the rise of rational expectations models in macroeconomics in the 1970s and the development of VAR models as well as DSGE models in the 1980s.

6.1.1 The VAR Representation of Traditional DSEMs

There is a close connection between VAR models and the structural, reduced-form and final-form representations of DSEMs. Zellner and Palm (1974) show that the reduced form of structural DSEMs in general has a VARMA (or VARMAX) representation that can typically be approximated by a finite-order VAR (or VARX) model.

We can think of the structural representation of an econometric model as the data generating process underlying the observed economic data. For an econometric model to be structural it is necessary for the stochastic shocks (or errors) in each equation to be mutually uncorrelated. This feature allows us to consider thought experiments in which one structural shock moves while leaving all other shocks unchanged. The resulting responses of the observables then represent the causal effects of this structural shock. This interpretation would break down if the shock in question were correlated with other shocks.

To illustrate the relationship between structural DSEMs and VAR models, consider the very stylized example of a structural model involving only two equations: (1) a function for aggregate consumption (c_t), built on the assumption of adaptive expectations, and (2) an equation for national income (n_t):

$$\begin{aligned} c_t &= \eta_1 + \alpha n_t + \beta c_{t-1} + w_{1t}, \\ n_t &= \eta_2 + \gamma c_{t-1} + \delta n_{t-1} + w_{2t}, \end{aligned}$$

where w_{1t} and w_{2t} are mutually uncorrelated white noise structural shocks. Any such structural model has a reduced-form representation, which may be constructed by rewriting the system until only lagged dependent variables are left on the right-hand side (RHS). Here it suffices to substitute the complete RHS of the second equation for n_t in the first equation. We obtain

$$\begin{aligned} c_t &= (\eta_1 + \alpha\eta_2) + (\beta + \alpha\gamma)c_{t-1} + \alpha\delta n_{t-1} + (w_{1t} + \alpha w_{2t}), \\ n_t &= \eta_2 + \gamma c_{t-1} + \delta n_{t-1} + w_{2t}. \end{aligned}$$

This reduced form can be expressed as

$$\begin{aligned} c_t &= \nu_1 + a_{11,1} c_{t-1} + a_{12,1} n_{t-1} + u_{1t}, \\ n_t &= \nu_2 + a_{21,1} c_{t-1} + a_{22,1} n_{t-1} + u_{2t}, \end{aligned}$$

where, for example, in the first equation $\nu_1 = (\eta_1 + \alpha\eta_2)$, $a_{11,1} = (\beta + \alpha\gamma)$, $a_{12,1} = \alpha\delta$, $u_{1t} = (w_{1t} + \alpha w_{2t})$. Equivalently, this model may be written in matrix notation as a first-order VAR(1) model:

$$y_t = \nu + A_1 y_{t-1} + u_t,$$

where $y_t \equiv (c_t, n_t)'$, $\nu \equiv (\nu_1, \nu_2)'$, $u_t \equiv (u_{1t}, u_{2t})'$ is white noise with mean vector 0 and covariance matrix Σ_u, and

$$A_1 \equiv \begin{bmatrix} a_{11,1} & a_{12,1} \\ a_{21,1} & a_{22,1} \end{bmatrix}.$$

This example illustrates that the parameters of a reduced-form model in general are combinations of the parameters of the underlying structural model. Econometricians seek to learn about the parameters of the structural model from the parameters of the reduced-form model. If the structural parameters can be recovered uniquely from the reduced-form parameters in population, we say that the structural parameters are exactly identified, point identified, or just identified. Determining the values of the structural parameters from the reduced form involves solving a system of equations. In the example above, this derivation is straightforward because the structure of the model is recursive: $\delta = a_{22,1}$, $\gamma = a_{21,1}$, and $\eta_2 = \nu_2$. Hence, $\alpha = a_{12,1}/a_{22,1}$, $\eta_1 = \nu_1 - \nu_2 a_{12,1}/a_{22,1}$, $\beta = a_{11,1} - a_{21,1}a_{12,1}/a_{22,1}$, $w_{2t} = u_{2t}$, and $w_{1t} = u_{1t} - u_{2t}a_{12,1}/a_{22,1}$.

In general, it is possible that some structural model parameters remain unidentified (which means that there are infinitely many possible solutions for these structural parameters even in a sample of infinite length) or overidentified (which means that there are more restrictions than needed to recover the structural parameters). Models in which not all structural parameters are exactly identified are called underidentified or partially identified. Finally, in the special case of models identified based on inequality restrictions, the structural parameters will be set-identified (which means that we can only bound the set of structural parameter values). We defer discussion of such models to Chapter 13.

With these definitions in mind, let us return to the previous example. Clearly, if we had written down a different structural model for c_t and n_t, we might have obtained a different reduced-form representation. Regardless of which structural model is the data generating process for c_t and n_t, however, if we had specified a reduced-form VAR(p) model for $y_t \equiv (c_t, n_t)'$ including a sufficient number of autoregressive lags such that

$$y_t = \nu + A_1 y_{t-1} + \cdots + A_p y_{t-p} + u_t,$$

we would expect that reduced-form model to be logically consistent with a large set of different structural models. This intuition also applies to higher-dimensional models. In short, reduced-form VAR models encompass a range of structural DSEMs. This stylized example could be generalized to allow for the inclusion of exogenous variables. We do not consider this extension because exogenous variables are rare in the VAR literature for reasons discussed in the next subsection.

6.1.2 *Incredible Restrictions in Traditional DSEMs*

As we have seen, the vector autoregressive model is closely related to the reduced-form representation of traditional structural DSEMs. The key difference is that the interpretation of the vector autoregression as a structural DSEM requires additional cross-equation restrictions and possibly exclusion restrictions on the reduced-form VAR parameters. These restrictions may also affect the dynamic structure of the model. There are two types of restrictions on the dynamic specification of the model. One involves treating some model variables as exogenous, consistent with the partial equilibrium nature of the IS-LM models of the 1950s and 1960s. This allows one to greatly reduce the number of parameters to be estimated.

Loosely speaking, a variable is exogenous if it is determined outside the system of equations under consideration. Such a variable is not subject to current or lagged feedback from other model variables, but only depends on its own lags (or on other exogenous variables). In practice, DSEMs treat a wide range of variables as exogenous with respect to the model variables of interest. Common examples include population growth, the money stock, and global commodity prices. More often than not, regarding a variable as exogenous is at best an approximation and at worst a heroic assumption. For example, the money stock is affected by the endogenous money creation of the banking system and hence not a plausibly exogenous model variable (see Chapter 7).

The other type of restriction involves constraining the dynamics of the endogenous variables and of the error term. Ideally, these dynamic restrictions should reflect economic theory, but traditional macroeconomic models such as the IS-LM model are static models. They offer no guidance for the specification of the dynamics of the econometric model. Such dynamics are important for two reasons. One reason is that adjustments to shocks are inherently sluggish, reflecting frictions in the economy. The other reason is that economic decisions depend on expectations which are driven at least in part by lagged model variables. In practice, therefore, econometricians in the 1950s and 1960s augmented the original static IS-LM model by adding dynamics, for example, in the form of partial adjustment models of investment, adaptive expectations models of consumption, and distributed-lag models for the regression errors or for exogenous model variables. The distributed-lag models often were restricted to enforce a smooth decay of the model coefficients at higher lags and to reduce the number of model parameters to be estimated freely. The purpose of these restrictions was (a) to enable researchers to fit large-dimensional models to the data and (b) to allow them to recover the structural parameters of the model. For further details see, e.g., Judge, Griffiths, Hill, Lütkepohl, and Lee (1985).

There are three problems with this traditional approach. First, commonly used dynamic specifications tend to violate the notion of rational expectations (which implies that economic agents do not make mistakes in expectation). This is a particular concern for partial adjustment and adaptive expectations specifications that prevent agents from adjusting fully on impact to policy changes. Second, the choice between different dynamic specifications often is arbitrary. Third, the building blocks of large-scale DSEMs are developed in isolation, with researchers focusing on one block of the model at a time (e.g., the consumption block, investment block, trade block, fiscal block, or monetary block), taking as given all model variables not determined within their block. This partial-equilibrium approach ignores the feedback from one model block to another in general equilibrium. Dynamic general equilibrium macroeconomic models imply that every variable depends on every other variable in the economy, which contradicts the traditional notion that some variables may be treated as exogenous with respect to others.

6.1.3 Structural VAR Models as an Alternative to Traditional DSEMs

Since the restrictions on lagged model variables in the DSEM are rarely credible and since the precise form of the structural data generating process is not known in practice, Sims (1980a) made the case that we might as well replace the restricted reduced-form model by an unrestricted VAR model with a suitably chosen maximum lag order. This approach, of course, forces us to reduce the dimensionality of the model greatly because a VAR model involves many more parameters to be estimated. Instead of considering hundreds of equations, we will be able to include only a handful of equations in the model. Moreover, without an alternative set of identifying assumptions, VAR models merely provide reduced-form summaries of the data that are not particularly interesting from an economic point of view. Sims' key insight was that, starting with a general structural model of the form

$$B_0 y_t = B_1 y_{t-1} + \cdots + B_p y_{t-p} + w_t,$$

with reduced-form representation

$$y_t = \underbrace{B_0^{-1} B_1}_{A_1} y_{t-1} + \cdots + \underbrace{B_0^{-1} B_p}_{A_p} y_{t-p} + \underbrace{B_0^{-1} w_t}_{u_t},$$

after setting $\Sigma_w = I_K$ without loss of generality such that $B_0^{-1} B_0^{-1\prime} = \Sigma_u$, we can dispense with ad hoc restrictions on the slope parameters altogether, provided we are able to restrict enough elements of the structural impact multiplier matrix B_0^{-1}, so the remaining elements can be estimated from the data. Put differently, knowledge of B_0^{-1} suffices to trace out the dynamic response of the model variables to structural shocks even in the absence of restrictions on the slope parameters. Econometricians soon developed a range of methods of imposing such identifying restrictions on B_0^{-1} (or, equivalently, on B_0). We review this structural VAR literature in Chapters 8, 10, 13, 14, and 15. To the extent that these restrictions are more credible economically than the restrictions in traditional DSEMs, the structural VAR approach provides a superior alternative to traditional DSEMs.

It is important to reiterate that Sims' critique was not directed at the idea of DSEMs itself, but rather at how these models were specified in practice in the 1970s. In fact, structural VAR models of the type studied in this book may be interpreted as small-scale DSEMs in which restrictions have been imposed on the impact multiplier matrix and the lagged values are left unrestricted.

6.2 The Relationship between VAR Models and DSGE Models

VAR models are also closely related to DSGE models. DSGE models require the user to explicitly specify the microstructure of the economy. The premise is that this model structure is invariant to policy interventions. Agents maximize their objective functions in expectation subject to individual and aggregate resource constraints, while taking account of the behavior of other agents in a rational manner. They do so in a world subject to random variation in variables such as technology that are considered exogenous with respect to agents' decisions in that there is no contemporaneous or lagged feedback from the agents' actions to these variables. Changes in "technology" here refer broadly to changes in the efficiency with which firms combine the factors of production into final goods. After log-linearizing the DSGE model about the nonstochastic steady-state path of the model variables, agents' behavior and choices can be characterized in the form of log-linear decision rules that depend only on state variables known to the agents at each point in time. This log-linear structure allows us to simulate the dynamic behavior of the endogenous model variables subject to the arrival of random shocks.

The state-space representation of DSGE models, under suitable conditions, can be written as a reduced-form VARMA model for the endogenous model variables, which in turn, under suitable conditions, can be approximated by a finite-order reduced-form VAR model. The link between structural VAR models and DSGE models is less straightforward and depends on how the structural VAR shocks are identified. For a more detailed discussion of how to derive the approximate VAR representation of DSGE models the reader is referred to the survey by Giacomini (2013). For further discussion of the link between DSGE and VAR models, see Christiano, Eichenbaum, and Vigfusson (2006b, section 2.2), Fernández-Villaverde, Rubio-Ramírez, Sargent, and Watson (2007), Ravenna (2007), Franchi and Vidotto (2013), and Guerron-Quintana, Inoue, and Kilian (2017). Next, we sketch the central issues in approximating DSGE models by finite-order VAR models and provide some intuition. Our discussion builds on Christiano (2002).

6.2.1 Basics

Let z_t be an $n \times 1$ vector of endogenous DSGE model variables (e.g., capital stock, output, consumption, investment). Let s_t be an $m \times 1$ vector of exogenous state variables (e.g., technology shocks, government spending shocks, money supply shocks, preference shocks). Without loss of generality, the law of motion of s_t can be expressed as a VAR(1) process

$$s_t = Ps_{t-1} + \varepsilon_t,$$

where the $m \times 1$ vector ε_t is zero-mean white noise such that ε_t is uncorrelated with ε_{t-l} and s_{t-l} for $l > 0$.

This law of motion is quite general and accommodates an arbitrary ARMA(p, q) representation for the underlying shocks. Note that the common terminology of referring to s_t as "shocks" is misleading from an econometric point of view, because by construction the shock is ε_t rather than s_t. In reality, s_t refers to an exogenous state variable in the model. We nevertheless will follow the DSGE model literature in using this terminology in this section.

The leading example for an exogenous state variable is the technology shock process underlying the neoclassical growth model. We denote this variable by x_t. It is usually assumed that the process for technology (which determines aggregate productivity levels in the model economy) is highly persistent and follows an AR or ARMA process.

Example 1 Suppose a shock x_t follows an AR(1) process

$$x_t = \rho x_{t-1} + e_t,$$

where e_t is mean zero white noise. Then $s_t = Ps_{t-1} + \varepsilon_t$, with

$$s_t = x_t, \quad P = \rho, \quad \varepsilon_t = e_t.$$

Example 2 If x_t follows an ARMA(1,1) process

$$x_t = \rho x_{t-1} + e_t - \gamma e_{t-1},$$

instead, then $s_t = Ps_{t-1} + \varepsilon_t$, with

$$s_t = \begin{pmatrix} x_t \\ e_t \end{pmatrix}, \quad P = \begin{bmatrix} \rho & -\gamma \\ 0 & 0 \end{bmatrix}, \quad \varepsilon_t = \begin{pmatrix} e_t \\ e_t \end{pmatrix}.$$

It can be shown that a broad class of models, including real business cycle models with taxes and other distortions, limited participation models, models of labor hoarding, and models with sticky prices, has solutions for the endogenous model variables of the form:

$$z_t = Az_{t-1} + Ms_t,$$

where A is $n \times n$, M is $n \times m$, and $s_t = Ps_{t-1} + \varepsilon_t$. Hence, the reduced form representation of the vector process z_t will not in general have a finite-order VAR representation, unless $P = 0$ and s_t reduces to white noise. In fact, the process z_t need not even have a finite-order VARMA representation.

Economic theory provides no guidance for the specification of the time series process of the exogenous state variable s_t, but it is generally agreed that s_t must be fairly persistent for models to generate persistence in z_t comparable to the actual data. Hence, s_t will not be white noise. Typical choices include AR(1), AR(2), or MA(2) models for technology, the growth in monetary aggregates and in government spending, and other variables treated as exogenous state variables.

The specification of the latent state variables, s_t, matters for the time series representation of the endogenous model variables, z_t. For example, choosing a process for the technology shock, x_t, has immediate implications for the time series representation of endogenous model variables such as the capital stock. In the neoclassical growth model, for example, the household's capital stock for next period, k_t', is decided one period in advance and evolves according to the log-linear decision rule:

$$k_t' = \alpha_0 + \alpha_1 k_{t-1}' + \alpha_2 x_t, \tag{6.2.1}$$

where x_t denotes the technology shock. Thus, if x_t follows an MA(1) process, k_t' will follow an ARMA(1,1) process. If x_t follows an AR(1) process of the form $(1 - \rho L)x_t = e_t$, however, as is commonly assumed in the literature, k_t' will follow an AR(2) process. This result can be verified by premultiplying (6.2.1) by $(1 - \rho L)$.

The variable k_t' is, of course, only one of many endogenous variables in the DSGE model. In practice, we are interested in the joint n-dimensional vector time series process z_t, which raises additional complications. It may seem that z_t would follow a VARMA process if one of its elements follows an ARMA process, but this is not necessarily the case. A common problem is that the number of exogenous shocks in the DSGE model is smaller than the dimensionality of z_t. In this case, the time series process for z_t will be of reduced rank. For example, a textbook real business cycle (RBC) model has only one shock, so, when fitting a VAR(MA) model to output, investment, and consumption data from this RBC model, the error covariance matrix Σ_u would be of reduced rank. It follows immediately that z_t cannot be approximated by conventional VAR(MA) processes because the latter model requires the error covariance matrix Σ_u to be of full column rank. This condition means that there must be (at least) as many shocks as variables in the VAR(MA) model.

This rank-deficiency problem can be remedied in four ways: (a) by the use of reduced-rank estimation methods; (b) by adding noise shocks to the DSGE model (often referred to as measurement errors, although model approximation errors would be a more apt name) that are devoid of any economic interpretation, thereby undoing the benefits of relying on explicit microfoundations; (c) by reducing the number of observables in the VAR model; or (d) by augmenting the number of economically interpretable shocks in the model. Of course, this raises the question of why the latter shocks were not included in the model in the first place. One common response is to include preference shocks and efficiency shocks in the DSGE model that allow for departures from first-order conditions that must hold in equilibrium. The difference between such shocks and noise shocks is largely a matter of degree. Neither type of shock is based on microfoundations. For now, let us suppose that the reduced-form representation of z_t is of full rank.

6.2.2 The Role of Data Transformations

In practice, it is almost never the case that we fit a parametric VAR model to the n-dimensional vector z_t. One reason is that there are often more variables in the DSGE model than could be reasonably included in a small-scale to medium-scale VAR(MA) model. Hence, we inevitably must integrate out one or more of the elements of z_t. Another reason is that not all time series in the model are actually observed by the econometrician. For example, there are no good data on the capital stock that could be included in a VAR model, which again requires the marginalization of z_t.

This marginalization affects the time series process for the remaining endogenous model variables, as discussed in Chapter 2. In particular, even if z_t followed a known parametric VAR(p) process, any K-dimensional subprocess or marginal process y_t of an n-dimensional process z_t, obtained by using transformation matrices of the form $F = [I_K, 0]$ with $K < n$, would not in general be a VAR process of finite order. This result follows from the following proposition in Lütkepohl (2005):

Proposition. Let z_t be an n-dimensional, stable, invertible VAR(p) process and let F be a $K \times n$ matrix of rank K. Then the process $y_t = Fz_t$ has a VARMA(p^*, q^*) representation with $p^* \leq np$ and $q^* \leq (n-1)p$. □

Hence, except in rare cases, DSGE models will not have a finite-order VAR representation. Moreover, even if the data from a particular DSGE model were to have a finite-order VAR representation, any aggregation of the observed data not accounted for in the DSGE model would similarly induce MA components of unknown form in the reduced-form representation of the data. One example is statistical agencies aggregating data across households or firms. Another example is the aggregation of data over time (say, from monthly to quarterly frequency). Thus, it would be surprising indeed if the quarterly VAR representation of y_t were truly of finite lag order (see also Section 2.2.3).

6.2.3 Why Not Use VARMA Models?

In light of these arguments against finite-order VAR models, it may seem that we simply should have estimated finite-order VARMA(p, q) models instead. There are two counterarguments, however. First, we have little confidence in the particular VARMA(p, q) specification implied by a given DSGE model because that specification depends on inherently atheoretical assumptions about the dynamics of the exogenous state variables. Second, large-dimensional VARMA models are difficult to estimate reliably and hence are rarely used.

These considerations suggest the need for an alternative class of reduced-form models when approximating data generated by DSGE models. If the VARMA representation is invertible, z_t can be represented as a VAR process. More often than not, this VAR process is of infinite order. In this case, we can appeal to standard results in the literature about approximating VAR(∞) processes.

6.2.4 Autoregressive Sieve Approximations of VAR(∞) Processes

Assuming an exponential rate of decay of the coefficients of the autoregressive representation, the VAR(∞) process may be approximated by fitting a sequence of finite-order VAR(p_T) models, as discussed in Chapter 2. The approximation error becomes arbitrarily small asymptotically, provided that p_T increases with the sample size T at a suitable rate. This semiparametric approach is also known as an autoregressive sieve approximation. The sieve idea refers to the fact that the choice of p_T determines how much information passes through the sieve. This sieve approach is not a panacea, however. First, it will only be suitable if the DSGE model is invertible. Second, a good linear approximation for given T may require a very large p_T.

One concern in the literature is that in some cases for small T no feasible choice of p_T may suffice for a good approximation of the structural impulse responses implied by a DSGE model. Given the difficulty of matching structural shocks in VAR models with structural shocks in DSGE models, the literature addressing this point has tended to focus on VAR models of the effects of technology shocks identified by long-run identifying restrictions (see Chapter 10). There are examples in which structural impulse response estimates, generated by fitting such VAR models to data generated from DSGE models, do not even come close to recovering the structural impulse responses implied by the underlying DGP. For example, Ravenna (2007) finds that even VAR models with 12 lags estimated on 50 years of quarterly data may provide a poor approximation to the structural responses of interest (see also Mertens 2012; Liu and Konstantinos 2012; Poskitt and Yao 2017).

Kascha and Mertens (2009) conclude that this problem is not so much related to the quality of the VAR approximation as to the use of long-run restrictions for identification (see Chapters 10 and 11). They show that VARMA models are no more accurate than VAR approximations as measured by the bias and MSE of the structural impulse responses. One possible explanation is that the VARMA model is close to being unidentified because the roots of the AR and MA polynomials are similar in magnitude, nearly causing these roots to cancel (see Andrews and Cheng 2012). However, Kascha and Mertens find that equally inaccurate estimates are also obtained when estimating the reduced-form MA representation directly from the state-space equations of the DSGE model, which eliminates the problem of model misspecification and the problem of near root cancellations. How general this type of result is remains an open question. In related work, Pagan and Robinson (2016) provide evidence that the quality of low-order reduced-form VAR approximations to a DSGE model also depends on whether the stock variables in DSGE models (such as the capital stock), which are typically not observed in the real world, can be expressed as a function of a small number of lags of the observed variables.

A final concern is that none of these studies makes allowances for the small-sample bias in the VAR slope parameter estimator $\widehat{A}$ discussed in Chapter 2. This small-sample bias arises even in correctly specified models and may greatly undermine the accuracy of structural impulse response estimators, especially when working with model variables expressed in log-levels (see, e.g., Kilian 1998c).

6.2.5 Summary of Potential Problems in Approximating DSGE Models with VAR Models

To reiterate, there are several conditions required for approximating the reduced-form representation of DSGE models with VAR models. First, the number of shocks in the DSGE model must match the number of shocks in the reduced-form VAR model. We implicitly assumed that this condition is met, but this condition must be checked on a case-by-case basis. Second, the state-space representation of the DSGE model must be invertible for a VAR approximation to work. Fernández-Villaverde, Rubio-Ramírez, Sargent, and Watson (2007) study conditions under which the state-space representation of DSGE models and VAR models match up. They show that invertibility cannot be taken for granted. This question is closely related to the observation going back to the early 1980s that models with forward-looking agents may not have a (backward-looking) VAR representation (see also Chapter 17). Despite these caveats, Fernández-Villaverde et al. conclude that many models of interest to macroeconomists are invertible. Moreover, Sims (2012) shows that, even when the conditions for the invertibility of the state-space representation fail, the degree of misspecification of the structural VAR responses may be small.

Even if these first two conditions hold, however, the question remains of how good of an approximation finite-dimensional VARs provide to the VAR(∞) data generating process. Simulation evidence shows that in some examples the VAR(p_T) approximation to the VAR(∞) model appears adequate. In other examples, it may be poor for any feasible choice of p_T in realistic sample sizes, which may be an indication of the process being nearly noninvertible, but also could be caused by other approximation errors. How important such counterexamples are depends on how relevant the latter DSGE models are for policy work. An obvious question for future research is how good the autoregressive approximation is for DSGE models commonly used by policymakers. Another important question is how to select the approximating lag order p_T in practice. The answer to this question also depends on which aspect of the fit of the DSGE model we are interested in. This is an open area of research. For example, Hall, Inoue, Nason, and Rossi (2012) propose a procedure for selecting p_T in conjunction with the horizon of the impulse responses to be matched.

All of the potential problems discussed so far relate to the question of whether reduced-form VAR models can approximate data generated by DSGE models. If we want to compare impulse responses from a structural VAR model to impulse responses in DSGE models, we must make sure in addition that the identifying assumptions in the VAR are fully consistent with the structure imposed in the DSGE model. This requirement is difficult to meet in practice. For example, the timing assumptions imposed in standard semistructural VAR models of monetary policy, as discussed in Chapter 8, are at odds with the structure of many DSGE models of monetary policy. As a result, caution must be exercised in comparing results from DSGE and structural VAR models. This caveat is likely to apply more to some methods of achieving identification in structural VAR models than to others. For example, structural models based on sign restrictions can be expected to be more robust than methods based on short-run exclusion restrictions. We will return to this discussion in Chapter 11.

6.3 DSGE Models as an Alternative to VAR Models?

The force of the Lucas critique of DSEMs led to considerable interest in rational expectations models in the 1970s and early 1980s. This movement was led by Thomas Sargent and Robert Lucas. Whereas there is little controversy among economists that agents are rational in the sense of not making systematic mistakes, that idea is devoid of empirical content if we do not know the objectives and constraints of agents. In practice, this fact gave rise to the Rational Expectations Hypothesis which states that agents' expectations must be consistent with the structure of the economic model, which also requires each agent to know and understand this structure. Thus, the question of testing rationality became intertwined with the question of testing the economist's ideas of what agents' preferences and constraints are. In other words, a rejection of this Rational Expectations Hypothesis may occur because agents are not rational or because the economist's model is wrong.

Early efforts to estimate small-scale rational expectations models were abandoned because the restrictions implied by rational expectations were routinely rejected, and the resulting models performed worse than unrestricted models (see, e.g., Hansen and Sargent 1980; Wallis 1980). Rational expectations made a comeback in the mid-1980s and 1990s, however, when medium-scale DSGE models became the dominant form of macroeconomic models used by academics.

6.3.1 Calibrated DSGE Models

Initially, DSGE models were calibrated rather than estimated. Users simulated the dynamics of the model variables conditional on a set of so-called deep parameters, chosen in a process referred to as calibration. Deep parameters here refer to parameters governing technology and preferences that are presumed invariant to policy changes. The central idea of calibration is to pin down as many deep parameters as possible based on the long-run mean of the data (the first moment), before informally assessing the fit of the model by comparing the cross-autocorrelations (the second moment) in the model and in the data. In practice, this approach is insufficient, however. Some of the remaining model parameters therefore are chosen based on extraneous microeconometric estimates. This approach is controversial because such estimates need not be representative for the behavior of the economy at the aggregate level. Finally, the persistence of the exogenous technology shock is chosen to match the persistence of real GDP in the U.S. data, violating the premise of not relying on second moments in calibrating the parameters (see, e.g., Prescott 1986; Cooley and Prescott 1995). The latter procedure can be viewed as an informal application of the method of moments subject to the remaining parameter values having been restricted. Once the process of calibration is complete, however, all calibrated model parameters are treated as known when confronting the model with the observed data.

Proponents of the calibration methodology dismiss conventional measures of model fit and the idea of statistical testing. Their argument is that all models are inherently misspecified, freeing us from the need to formally test the model. After all, a statistical rejection of the model would only confirm what they had already conceded from the start. Moreover, proponents of calibration prefer not to articulate a formal loss function to be used in evaluating the fit of the DSGE model, allowing them to decide which dimension of model fit is most important on a case-by-case basis.

When calibrators confront their model with the data, they do so informally by inspecting tables of cross-autocorrelations in the data and comparing them with the cross-autocorrelations in the DSGE model. If the model appears at odds with the data in some dimension that is considered important, this fact is interpreted not as a rejection of the model, but as an indication that the model structure requires further refinements. The basic modeling paradigm is taken as correct by assumption. The approach of trying to match unconditional second moments in the model and in the data may look at first sight like an informal application of the method of moments, but it is not because – rather than adjusting the values of the parameters in the DSGE model in response to indications of poor fit – calibrators tweak the DSGE model structure to achieve improved fit, conditional on the same set of calibrated parameters.

6.3.2 Estimated DSGE Models

The calibration approach to evaluating DSGE models, as envisioned originally, has not stood the test of time. Instead, users of DSGE models at central banks and most academic researchers nowadays rely on direct estimates of the state-space representation of DSGE models. The development of such estimation methods started about 10 years after the publication of Kydland and Prescott (1982). This does not mean that calibration is not a useful tool for macroeconomists. It may be reinterpreted as a method for evaluating the quantitative implications of economic theory, the focus being on understanding the dynamic implications of the DSGE model, not on fitting the actual data. If we are interested in the fit of DSGE models and in their ability to explain observed data, in contrast, there is no substitute for econometric methods. The first study to attempt to estimate the structural representation of a DSGE model directly by unrestricted ML estimation was Leeper and Sims (1994). This research program has been largely abandoned because of convergence problems arising from the large dimension of the parameter space.

An alternative estimation approach in the DSGE model literature is the method of moments. Given that the second moments of the data implied by the DSGE model can be expressed as a function of the deep model parameters, one may infer the values of these deep parameters by matching the moments in the DSGE model with the corresponding moments in the data. Based on a simulation study, Ruge-Murcia (2007) concludes that generalized method of moments (GMM) estimators and simulated method of moments estimators based on an auxiliary time series model compare favorably with the unrestricted ML estimator of the DSGE model. Another popular estimation method is impulse-response matching (see, e.g., Guerron-Quintana, Inoue, and Kilian 2017).

Although there continues to be interest in frequentist estimation methods for DSGE models, the dominant estimation approach since the 2000s has been Bayesian. For this purpose, the structural state-space representation of the DSGE model is typically written in the form of a Gaussian likelihood and combined with prior distributions for the structural DSGE model parameters. One reason for the appeal of the Bayesian framework is that Prescott's calibration method can be viewed as Bayesian inference with a degenerate prior, conditional on the data, so full-fledged Bayesian methods of inference with proper priors on the structural parameters are a natural generalization of his ideas (see Del Negro and Schorfheide 2008). Another reason is that Bayesian priors can be used to restrict the search path of numerical optimization routines, facilitating the convergence of the estimates to economically plausible values. For a review of Bayesian estimation methods for DSGE models, see An and Schorfheide (2007).

6.3.3 Calibration versus Bayesian Estimation

There are some interesting parallels between calibrators and Bayesian econometricians. The Bayesian approach has a long tradition in time series econometrics. As discussed in Chapter 5, Bayesians think of the data as fixed and the DSGE model parameters as uncertain according to a prior distribution with suitable support. In contrast, calibrators may be viewed as members of a wayward Bayesian tribe whose members think of the data as fixed but of the DSGE model parameters as having a degenerate prior distribution obtained by calibration. In other words, they claim to be absolutely certain of their prior views about the model parameters before looking at the data.

This means that in principle they could simply work out the prior mean of the finite-sample distribution of the statistics of interest in the DSGE model (say, a coefficient of the cross-autocorrelation function of the model data) and compare it to the value of the same statistic in the actual data. Because in practice this finite-sample distribution is not known, calibrators tend to simulate the finite-sample distribution and to compute the average of the statistics of interest numerically. They then compare this prior mean to the corresponding statistic in the actual data.

Given that the prior distribution is degenerate, however, users of this methodology do not update their views about the model parameters based on the data, as one usually would in Bayesian analysis, by forming a posterior distribution. Indeed, calibrators would think of evidence of poor fit between the prior mean of the statistic of interest and the observed value of the same statistic in the data not as a reason to change the values of the model parameters, but at best as a reason for refining the DSGE model structure. This is a key difference between calibrators in the tradition of Kydland and Prescott and Bayesian econometricians in the tradition of Sims. Finally, the standard errors for the statistics of interest one sometimes sees reported in the literature on calibrated DSGE models do not relate to prior uncertainty about the parameters (because there is none), nor do they relate to the posterior uncertainty about the parameters (because no estimation is involved in their construction). Hence, they cannot be used to construct regions of highest posterior density or to construct confidence intervals in the frequentist sense.

6.3.4 Are Structural VAR Models Less Credible than DSGE Models?

A common misperception among macroeconomists in recent years has been that DSGE models are "structural", whereas structural VAR models are not. The unspoken assertion is that estimates of structural VAR models are at least debatable, if not flawed, and inherently inferior to estimates of DSGE models. This confusion arises because the term "structure" referred to by these macroeconomists differs from the way this term is used by econometricians. In econometrics, the term "structural model" refers to the conditions required for identifying structural shocks, as formulated by the Cowles Commission in the early 1950s.[1] Among DSGE modelers it refers to explicit assumptions about the market microstructure in the economy, about agents' constraints, and about the functional form of their objectives.

This DSGE microstructure is not required for the identification of structural shocks. Its potential benefit is that it implies additional cross-equation restrictions on the VAR model, which may enhance the efficiency of the model estimates compared with a VAR model whose lagged coefficients remain unrestricted. These same restrictions, however, may also render the VAR model estimates inconsistent if the microstructure in the DSGE model differs from that in the actual economy. This concern is heightened by the observation that the misspecification of any part of the DSGE model in a general equilibrium setting also invalidates the estimates of parameters in other parts of the DSGE model.

Microfoundations refer to explicit assumptions about economic structure at the level of household and firm decisions, but these assumptions often are merely artificial devices intended to induce more realistic macroeconomic behavior. Examples include lotteries that determine whether a worker is unemployed or whether a firm is allowed to change prices, assumptions that prices or wages remain fixed for a predetermined period of time, the assumption that money by itself generates utility, or that households are subject to cash-in-advance constraints. Often these assumptions are not even remotely plausible, but reflect the need for tractable models or the limits to our modeling skills. The reliance on the representative-agent framework is one example. Another example is the nature of the assumptions about the functional form of aggregate production functions and utility functions. A third example is the ad hoc price and wage setting behavior in standard New Keynesian DSGE models. A fourth example is the simplistic structure of the labor market in DSGE models. Even more troublesome is that not only technology, but also monetary aggregates and government spending are treated as exogenous processes in many DSGE models. Likewise, the dynamic specification of the processes for the

[1] For a review of the Cowles Commission's approach to econometrics, see Christ (1994). Further discussion can be found in Cooley and Leroy (1985).

exogenous state variables is invariably ad hoc. In short, a model being rigorous in the sense of deriving results from explicit assumptions is not the same as that model being realistic. This does not mean that additional microstructure cannot be helpful in thinking about the economy, but that it is less than obvious that we would wish to impose this structure in estimating VAR models.

The same concern about model misspecification applies when estimating DSGE models directly as an alternative to the estimation of structural VAR models. This is not the only problem, however. There is also strong evidence that the structural parameters in the state-space representation of DSGE models are only weakly identified. Weak identification means that the likelihood is nearly flat across the parameter space, which implies that the posterior of the structural model parameters will be dominated by the prior. Put differently, whatever prior we impose in estimation is also for all practical purposes the posterior we obtain from the data. This observation tells us that we have learned nothing from the data. Evidence of weak identification casts doubt on standard frequentist and Bayesian approaches to estimating DSGE models. Econometricians are only beginning to develop methods of inference that are robust to weak identification problems in DSGE models (see, e.g., Guerron-Quintana, Inoue, and Kilian 2013; Dufour, Khalaf, and Kichian 2013; Qu 2014; Andrews and Mikusheva 2015). Many DSGE model estimates therefore have to be viewed with caution.

Proponents of DSGE models have at times made much of the fact that VAR models require auxiliary assumptions about the lag order and about the data transformations (see Cooley and Dwyer 1998). Because these auxiliary restrictions are atheoretical and effectively untestable, this argument goes, the resulting structural models are not as credible as DSGE models. It is easy to overlook, however, that much the same type of auxiliary assumptions are also required in specifying a DSGE model. For example, we may model the exogenous technology process in a DSGE model alternatively as a trend stationary process or as a process in differences, which is no different from deciding whether to express real GDP in a VAR model in growth rates or in deviations from a deterministic time trend. Moreover, this process may be specified as an AR(1) or an AR(2) process, for example, which is the same problem as choosing the lag order of the VAR model. None of these specification choices are grounded in economic theory.

Finally, an additional auxiliary assumption invoked by many users of DSGE models relates to how the simulated data should be filtered before the unconditional second moments can be computed (see also Chapter 19). DSGE models are designed explicitly as models of the business cycle and abstract from secular growth (see Kydland and Prescott 1982). Models of the business cycle should only be evaluated using data measured at business cycle frequencies, but observed macroeconomic data fluctuate at higher and lower frequencies, creating a mismatch between the model data and the actual data. For this reason, users of DSGE models often insist on transforming both the actual and the simulated data to isolate the variation at business cycle frequencies. Depending on the specification of the technology process, the choice is usually between deterministic detrending and HP-filtering the seasonally adjusted data or between differencing and HP-filtering, but other forms of filtering the data could also be entertained. This choice is not only ad hoc, but even seemingly reasonable approaches to detrending the simulated model data and the actual data may bias statistical measures of model fit and mask the propagation mechanism at work in the model, as illustrated in Singleton (1988), Diebold and Kilian (2001), and Canova (2014). This means that the "stylized facts" users of DSGE models aim to emulate with calibrated DSGE models are as likely to be artifacts of the filtering method as features of the data. The same concern applies when estimating the parameters of DSGE models on filtered data.

6.3.5 Are DSGE Models More Accurate than VAR Models?

Much of the literature on comparing DSGE models and VAR models has taken the DSGE model as given, while asking whether the VAR model can recover the features of the data generated by the DSGE model. The specification of the VAR model in this type of exercise is intended to reflect the specification of the DSGE model that is presumed to be the DGP. The rationale for this approach is by no means self-evident. As observed by Canova and Ciccarelli (2013, p. 206),

> tightly parameterized DSGE models are useful because they offer clear answers to policy questions and provide easy-to-understand welfare prescriptions. However, by construction, they impose a lot of restrictions that are not always consistent with the statistical properties of the data. Thus, the policy prescriptions are hardwired in the assumptions of the model, and must be considered more as a benchmark than a realistic assessment of the options and constraints faced by policymakers in real-world situations.

With equal justification one might ask whether the DSGE model is consistent with a VAR model chosen to fit the data. Reports in the literature about DSGE models fitting the data better than VAR or Bayesian VAR (BVAR) models hence have to be viewed with caution if the specification of the VAR model is artificially constrained to match the specification of the DSGE model.

There are three metrics for judging the relative accuracy of DSGE models and VAR models. One metric is in-sample fit measured, for example, by the marginal likelihood of these models (see Chapter 5). The marginal likelihood, unlike measures such as R^2, involves an implicit penalty for model complexity to avoid overparameterization (see Del Negro, Schorfheide, Smets, and Wouters 2007). Del Negro et al. report that the VAR specification subject to the cross-equation restrictions implied by the DSGE model fits the data better than the unrestricted VAR model. This conclusion, of course, applies to the VAR representation of this specific DSGE model only. No other VAR models are considered, so the result cannot be interpreted as this DSGE model being more accurate than the entire class of VAR models.

Another concern is that the priors used in Bayesian DSGE model estimation are heavily influenced by the data to be explained, violating the central premise of Bayesian analysis that the prior should not be influenced by the data (see Chapter 5). This problem arises because Bayesian DSGE models tend to rely on priors that have been found to work well in previous studies using nearly the same data set. Moreover, as the prior evolves more quickly than new data become available, there is a risk of converging to a posterior that fits the data well, perhaps even better than a VAR model, but is nevertheless spurious (see Kilian 2007).

One way of at least reducing this overfitting problem is to focus on the out-of-sample fit of the DSGE model. Until recently, the perception among proponents of DSGE models had been that DSGE models are the preferred structural models, but that reduced-form VAR models remain the benchmark when it comes to forecasting out of sample beyond the end of the estimation sample. Recent work by Del Negro and Schorfheide (2013) has challenged this view. Del Negro and Schorfheide presented evidence that Bayesian DSGE models are more accurate out-of-sample forecasting models than both unrestricted VAR models and BVAR models estimated in the Minnesota tradition. This point is important in that it suggests that the VAR model may no longer be the right benchmark for DSGE model evaluations.

Subsequent research has questioned this assessment. In particular, Gürkaynak, Kisacikoglu, and Rossi (2013) demonstrate that the relatively large-scale VAR and BVAR forecasting models considered by Del Negro and Schorfheide are not a credible forecasting benchmark because they are frequently outperformed by lower-dimensional autoregressive models. Gürkaynak et al. compare Bayesian DSGE models with a range of AR, VAR, and BVAR forecasting models for inflation, growth, and interest rates. They find that in real time no forecasting model is uniformly more accurate than the others. The relative performance of the models differs greatly over time, across forecast horizons, and across the variables to be forecast.

The third metric for comparing DSGE models and SVAR models is how close the implied structural impulse response functions are. There is in fact an entire literature on impulse response matching between DSGE models and VAR models based on minimum distance estimators (see, e.g., Rotemberg and Woodford 1997; Christiano, Eichenbaum, and Evans 2005; Dridi, Guay, and Renault 2007; Hall, Inoue, Nason, and Rossi 2012; Guerron-Quintana, Inoue, and Kilian 2017). A comparison of the structural impulse responses makes sense, however, if and only if the identification of the structural shocks in the VAR model is valid and is consistent with the structure imposed in the DSGE model. In practice, it can be difficult to find identification schemes that are consistent with the DSGE model and simultaneously identify an entire vector of structural shocks in a high-dimensional VAR model.

Some researchers have addressed this problem by imposing structure on the impact multiplier matrix of the VAR that comes from the fitted DSGE model (see Del Negro and Schorfheide 2004; Del Negro, Schorfheide, Smets, and Wouters 2007). This approach bypasses the issue of the identification of the structural VAR shocks, but requires a high degree of confidence in the DSGE model estimate. Moreover, Sims (2007) observes that the results of Del Negro and Schorfheide's procedure can be sensitive to the ordering of the VAR variables. Other researchers have resorted to lower-dimensional models and/or models that are only partially identified. A popular choice are models that identify technology shocks based on long-run identifying restrictions. Such models, however, are not without their own limitations, as explained in more detail in Chapters 10 and 11.

6.3.6 Policy Analysis in DSGE Models and SVAR Models

Another common misperception among macroeconomists is that structural VAR models cannot be used to conduct meaningful policy experiments. The central question in this debate is whether estimates of structural VAR models remain stable following government policy interventions. The Lucas critique suggests that only deep parameters (such as parameters representing tastes, preferences, and perhaps technological constraints) that govern the behavior of economic agents remain unchanged in response to policy interventions. This result seems to imply that reduced-form VAR representations cannot be expected to remain stable in response to policy interventions. This line of reasoning is frequently invoked as an argument (a) against relying on SVAR models and (b) for relying on DSGE models. Neither argument is compelling in general.

The first point to keep in mind is that Lucas' argument may apply to changes in policy regimes within the VAR framework, but typically does not apply to policy surprises within a given policy regime. As long as the policy surprises are within the range of historical experience, agents will have formed expectations that allow for these policy surprises and there is no reason for the reduced-form model to change (see, e.g., Sims 1986; Leeper and Zha 2003). To the extent that policy counterfactuals and forecast scenarios based on structural VAR models can be mapped into sequences of structural shocks, as discussed in Chapter 4, this reasoning may be applied even to changes in policy regimes.

 https://doi.org/10.1017/9781108164818.007

Some changes in the policy regime may be too dramatic to avoid the Lucas critique of reduced-form VAR models, but others will involve policy shocks that are well within the range of historical experience and are not predictable. The analysis of the latter policy regime shifts will not be in obvious conflict with the Lucas critique.

A second implication of the Lucas critique is that the analysis of monetary policy VAR models, in particular, ought to focus on samples during which the policy reaction function remained stable. Fitting monetary policy VAR models on data from the early 1970s until today, for example, cannot be expected to generate meaningful results. This problem may be addressed by splitting the sample. Limiting the sample to homogenous periods, of course, reduces the sample available for estimation and inference. Another possible response to this problem therefore is the use of time-varying coefficient VAR models or of Markov-switching VAR models that allow the policy regime to change over time, as discussed in Chapters 18 and 19 (see, e.g., Sims and Zha 2006b; Sims, Waggoner, and Zha 2008).

It is perhaps less widely appreciated that estimates of micro-founded DSGE models are not immune from the Lucas critique either. A case in point is the invalidity of estimates of time-invariant DSGE models in the presence of shifts in monetary policy regimes. For example, fitting a DSGE model with a standard Taylor policy rule on data ending in 2014 without accounting for the shift of monetary policy toward quantitative easing in 2008 would not produce meaningful estimates. Any misspecification of the monetary policy rule in the estimated DSGE models invalidates the estimates of all model parameters, just like in VAR models.

The sense in which a DSGE model may in principle be used to address the Lucas critique is that estimates of the deep parameters of a DSGE model subject to one policy regime (obtained from a sample devoid of any policy regime shifts) may be used to predict agents' behavior in a different policy regime. The prediction may be constructed by evaluating the DSGE model that incorporates the alternative policy regime conditional on the estimates of the deep parameters from the estimated model. This reasoning is based on the premise that deep parameters are invariant to changes in the policy regime and hence not subject to the Lucas critique. This premise is not uncontroversial (see Fernández-Villaverde and Rubio-Ramírez 2007). It also relies on the model structures in question being an accurate representation of reality before and after the policy regime shift. Finally, it hinges on the researcher's ability to estimate the deep parameters reliably. Evidence that deep parameters in commonly used DSGE models are only weakly identified casts doubt on the latter premise (see, e.g., Canova and Sala 2009).

We conclude that mechanical applications of either approach should be avoided. DSGE models and structural VAR models are complementary, with each approach having its own strengths and weaknesses. There is no basis for claims that one approach dominates the other.

6.4 An Overview of Alternative Structural Macroeconometric Models

The following table compares the three main approaches to empirical macroeconomics in the literature: the DSGE model, the DSEM, and the structural VAR (SVAR) model.

Features	DSGE	DSEM	SVAR
Exogeneity restrictions	Few	Many	None
Dynamic exclusion restrictions	Few	Many	Few
Number of variables	Large	Very large	Small
Number of shocks	Few	Many	Few
Trend treatment	Explicit	Implicit	Explicit
Microstructure required	Yes	No	No

The table highlights that each approach has its own strengths and weaknesses. For example, traditional DSEMs require many unrealistic exogeneity restrictions, whereas SVAR models require no exogeneity assumptions at all. DSGE models, on the other hand, require some unrealistic exogeneity restrictions. Likewise, DSEMs require many dynamic exclusion restrictions, whereas SVAR models are unrestricted except for an upper bound on the lag order. DSGE models in turn do not directly restrict the dynamics of the endogenous model variables, but restrict the dynamics of the underlying exogenous state variables.

While DSEMs allow the inclusion of a large number of variables, standard VAR models and to a lesser extent DSGE models are much more limited in scope. They also have to make do with a smaller number of shocks. Both DSGE models and SVAR models explicitly account for the trending behavior of many macroeconomic time series, whereas in DSEMs the trends are often implicit. Finally, only DSGE models require explicit assumptions about the model's market microstructure including the specification of the deep parameters governing the preferences of households and production possibilities of firms, the choice of functional forms for utility and production functions, and the specification of market structures (e.g., perfect competition, monopolistic competition, imperfect competition).

It is not surprising in light of this comparison that there have been attempts to combine DSEMs and SVAR models as well as SVAR models and DSGE models in an effort to improve their performance.

6.4.1 Combining DSEMs and SVAR Models

In the 1980s and 1990s, many central banks experimented with modernized versions of DSEMs. Drawing on insights from time series analysis, these models paid more attention to the trend specification, relaxed the dynamic specification of the errors, embodied restrictions that preserve the long-run comovement of many variables, and reduced the number of variables and equations. These changes, however, also made it more difficult to interpret the estimates economically because of the uneasy coexistence of reduced-form and structural elements in the same model. These models appear to have been abandoned in favor of estimated DSGE models in recent years.

At the same time, VAR users remained cognizant of the fact that standard VAR models are not able to compete with DSEM models in providing detailed answers about the response of the economy to structural shocks. Leeper, Sims, and Zha (1996), among others, exploited advances in computational techniques to build SVAR models with up to 20 variables. They relied on informative Bayesian priors to deal with the curse of dimensionality (see Chapter 5). Nevertheless, they could not overcome skepticism about their models' ability to extract useful information from the data, and such medium-scale VAR models never gained traction in empirical macroeconomics. More recently, there have been other efforts to extend the VAR framework to even larger-dimensional data sets, however. Examples are large-scale Bayesian VAR models and factor augmented VAR (FAVAR) models. These developments are reviewed in Chapter 16.

6.4.2 Combining DSGE and SVAR Models

With the demise of DSEMs in academic research, the focus shifted in the direction of combining DSGE models and SVAR models. Assuming that a VAR approximation to the DSGE model exists, the key difference between VAR models and DSGE models is that the latter imply additional cross-equation restrictions on the parameters of the VAR model that standard VAR models ignore. Unrestricted VAR models have the advantage that they are potentially consistent with a whole class of DSGE models. Their disadvantage is that additional structure, if correct, can reduce the VAR estimation errors in small samples. Much of the literature on combining DSGE and SVAR models has focused on the potential benefits of imposing DSGE structure on VAR estimates.

Early efforts to estimate small-scale SVAR models subject to restrictions implied by rational expectations models were abandoned because these restrictions were routinely rejected by the data (see, e.g., Keating 1990). Subsequently, King, Plosser, Stock, and Watson (1991) argued for imposing on VAR models the cointegration structure implied by RBC models with random walk technology shocks, resulting in a vector error correction model. Interest in forging a closer relationship between VAR and DSGE models increased a decade later, after DSGE models had become larger and embedded enough shocks and frictions to generate time series that look like actual data. Del Negro and Schorfheide (2004) in particular proposed priors for the estimation of VAR models that were derived from a New Keynesian DSGE model (see also Chapter 5). The extent to which such priors improve the accuracy of VAR estimates continues to be debated.

More recently, skepticism toward the structure of DSGE models has given rise to less parametric approaches of combining DSGE models and VAR models. For example, Giacomini and Ragusa (2014) advocate forcing the VAR model to satisfy the nonlinear equilibrium conditions in the DSGE model by exponential tilting. This approach is designed to impose only a subset of equations from the DSGE model, allowing us to discount DSGE model equations with less theoretical content or that we have less faith in. Similarly, Canova and Paustian (2011) rely on results from a broad range of DSGE models to derive identifying restrictions on the signs of structural impulse responses. This approach is discussed in more detail in Chapter 13.