MFE Economics
Lecture 1: Intuition

1. DSGE Models

DSGE stands for 'dynamic stochastic general equilibrium' which is the dominant class of macro models used in academic and policy circles at the moment. It emphasizes the fact that the real world involves decisions today, influenced by the past and in anticipation of the future, via forward looking expectations held by economic agents (firms and households). The term stochastic simply means that agents face and acknowledge randomness over what is going to happen or even, if they don't possess all information, what is or has happened. Generally we will think of 'known unknowns' as the main type of randomness (you know the odds) but sometimes people (see my research for example - if you're struggling with insomnia) refer to 'unkonwn unknowns' where you may not even know the odds. The former is typically referred to as 'risk' and the latter, 'uncertainty'. We will focus exclusively on the former. GE emphasizes that there are interconnections, spillovers and multiple directions of causality connecting endogenous variables (variables explicitly determined within the economy - rather than being primitive 'shocks' or uninfluenced by the actions of agents - which we call 'exogenous') in a way that respects individual optimality (utility and profit maximization, say) and market clearing and aggregate/technological feasibility.[1]

DSGE models need to be 'solved' for an equilibrium in which they make predictions about what will happen in certain contexts, given the 'state'. It's effectively the minimum amount of information you need to figure out what;s going on. In our simple models, the state will feature exogenous shocks (random impulses that hit the economy - like weather, unexpected fluctuations in policy...) and may also feature *predetermined* endogenous variables such as what wealth you have stored away from previous periods, or what machinery/capital has already been built. In equilibrium, a solved DSGE model entails a set of functions that map from the state to the current value of the non-predetermined (or 'jump') endogenous variables (together with the specification of how the exogenous shocks evolve, such as via an AR(1)).

To find these 'policy functions' (mapping from the state to the jump variables) we need to solve a bunch of equations, like we did in high school but with different interpretation. The equations are derived from household and firm first order conditions, the specification of the shock processes, market clearing and technological feasibility conditions (like the production function). We need as many equations as there are 'unknowns' (endogenous jump variables) to be solved for.

---

[1]Some models, more complicated than in this course, may also require various belief or other consistency requirements.

2. Stylized Facts

This section is fairly clear in the notes so I refer you to those. You should be aware of what the HP filter is - one of many reasonably sensible ways of breaking fluctuations into trend and cycle. Although both can be modeled jointly, it is (too) common for macroeconomists to focus on growth or business cycles separately . You should - as economists - be aware of various important stylized facts about both - and also be aware that after a long period of stability some of these facts have shown signs of changing in recent decades (notably labor share and trend growth rate in the U.S.). You should also be aware of how variables co-move over the cycle - which ones are pro-cyclical (positively correlated with output cycle), counter-cyclical or acyclical. Ultimately, models are designed to be consistent with these and other stylized facts.

3. Vector Autoregressions

You do not need to be a VAR specialist to understand this section or why we cover it - you will (almost certainly) be taught this in detail at some point this year in your time series econometrics. I am not teaching you it - so if it's new to you just realize that I do not cover it in enough detail for you to grasp it fully - so don't feel inadequate! Take a look at any econometrics textbook for a more thorough explanation and there are various notes online on 'Structural VARs'. Also, as I mentioned, see the first chapter of the Walsh book.

The main points I am trying to get across are as follows. VARs are a flexible way of characterizing how a set of variables evolve (stochastically) over time, where the process for their evolution allows for various feedbacks and lags (variables are correlated with eachother over timein a particular way). You can think of one equation in the VAR as saying that variable y depends on various lags of itself and on lags of another variable, x, plus a random innovation, u_y. You can say the same for variable, x, so you have two equations with pretty much the same form but with different coefficients on the RHS variables (which in this case are lags of blth variables). The VAR, when written in matrix form is just stacking these equations.

You can estimate the equations individually with OLS regression (I assume you've covered that my now in econometrics) and you can get estimates of the coefficients in the equations, arranged in a matrix, A (or matrices if you have more than one lag) and the covariance matrix of $u_t(u_{y,t}, u_{x,t})'$, which is a vector of mean zero forecast errors, call it Σ.

Given A and Σ you can derive expressions for the expected value of the variables y and x at any horizon, as well as the uncertainty around them. If you simply want to forecast and/or characterize moments, then this is enough but the VAR is, at this point, in 'reduced form'. Essentially this comes down to a question of the interpretation of u_t (the vector of 'reduced form innovations' or, more intuitively, 'forecast errors'. You don't need to know what 'caused' the forecast errors, *per se*, to characterize their distribution and then (using the VAR structure) the distribution of future values of $s_t \equiv (y_t, x_t)$. To see this, consider the 1 step ahead forecast

$$s_{t+1} = As_t + u_{t+1} \tag{1}$$

then the mean of s_{t+1} conditional in info in t is As_t since u_{t+1} is zero mean (conditionally or

unconditionally as we assume it is an iid $N(\mathbf{0}, \Sigma)$ vector innovation for all t.[2] The conditional variance of s_{t+1} is Σ. Both A and Σ can be estimated (OLS and use the covariance matrix of residuals to obtain the estimate of Σ) so then we can obtain forecast distributions under them. Now imagine forecasting two periods in advance, rather than one

$$s_{t+2} = As_{t+1} + u_{t+2} = A(s_{t+1} = As_t + u_{t+1}) + u_{t+2} = A^2 s_t + Au_{t+1} + u_t \tag{2}$$

so the conditional mean and variance are $A^2 s_t$ and (using independence of u_t over time) $A'\Sigma A + \Sigma$. More generally, for k periods ahead they are $A^k s_t$ and $\sum_{j=0}^{k}(A^j)'\Sigma A^j$.

Great, but suppose we think of y_t as a measure of real activity (GDP growth, say) and x_t a measure of monetary policy (interest rate, say), then one might want to know what the effect of an exogenous monetary policy shock (stemming from an unexpected change in the way the Fed thinks about the world - because of a mistake, or change in personnel) that manifests as a deviation from systematic policy. Typically we can't interpret $u_{x,t}$ as this without further assumptions. $u_{x,t}$ is the unexpected move (forecast error) in the interest rate, relative to info in $t-1$ but it might be because there was a surprise coming from another part of the economy in t - maybe a demand (confidence?) shock that will stimulate output (via a consumption or investment channel?) and which the Fed learns about *and reacts to*, also in t, under its existing policy approach (there is no structural monetary shock). Here the surprise increase in x_t (a tightening) stems not from a policy shock, but from a shock elsewhere and which also shows up in $u_{\mathbf{y},t}$ (the correlation mentioned in the previous footnote).

The fact that each element of u_t could be a mishmash (linear combination) of 'structural' shocks, call them $e_{y,t}$ (a confidence shock) and $e_{x,t}$, call it a policy shock, means that we can't look at the estimated response of the economy to either element of u_t and give it a structural, causal interpretation.

One formulation of this sort of problem (see slide 29) is to consider that there exists an unknown B matrix such that[3]

$$u_t = Be_{y,t} \tag{3}$$

where the covariance matrix of e_t is diagonal with some variances on the diagonal and zeros off diagonal, say

$$Cov(e_t) \equiv \Omega = \begin{pmatrix} \sigma_{e,x}^2 & 0 \\ c & \sigma_{e,y}^2 \end{pmatrix} \tag{4}$$

---

[2]The bold font for 0 reflects that it is a vector of zeros, not a single scalar. Note also it is the vector of errors that is iid - not each forecast error within a given period - they can be correlated - in fact this is closely related to the lack of a structural interpretation, discussed below.

[3]This B may exist because the underlying 'true' model of the economy states that certain variables are contemporaneously determined in equilibrium (think about later lectures of certain endogenous variables being dependent on the same shocks that are parts of the state

where, naturally, we have

$$\Sigma \equiv \begin{pmatrix} \sigma_{u,x}^2 & \sigma_{u,xy} \\ \sigma_{u,xy} & \sigma_{u,y}^2 \end{pmatrix} \equiv Cov(u_t) = B'\Omega B \tag{5}$$

This issue at play is that because, conceivably, both structural shocks can influence both forecast errors, we need to make assumptions on B that amount to assumptions on the sequencing of impacts. This allows us then to pull the separate structural shocks out of the reduced form ones and, consequently, figure out how thee economy responds to each of these structurally interpretable shocks. Note also that the structural shocks - in contrast to the forecast errors - are assumed to be contemporaneously uncorrelated - otherwise their interpretation becomes very unclear (in what DSGE model would that make sense? not many/any).

In the notes we discuss a way of doing this which is to make assumptions about what reacts to what. If we know one variable doesn't respond immediately to one of the structural shocks and one does, then all of the surprise in the first shock must be coming from the other structural. In the slides we see that you can choose different orderings determining which off diagonal element of B to set equal to 0.

Once you do that you can say ($\theta = 0$ example from the notes - and see also the first homework question on VARs) that $u_{y,t} = e_{y,t}$ and then $u_{y,t}$ can be taken to be $e_{y,t}$ (the 'confidence' shock). Then we observe that we can extract $e_{x,t}$ as $e_{x,t} = u_{x,t} - \phi u_{y,t}$ - but we don't yet know ϕ. In fact, by using equation (5) we see that $\phi = \frac{\sigma_{u,x}^2}{\sigma_{u,xy}}$ under our identification assumption (the restriction on B allows us to do this).[4] So if you've estimated the reduced form and have estimates for Σ, you can use those covariances - together with the identification assumption - to get at the structural covariance matrix and the underlying structural shocks can be obtained from your residuals, that are in a sense, estimates of the realizations of u_t.

Ultimately, the point of all this is for you to realize that people often use VARs to estimate the effect of monetary policy - and thus need some sort of identifying assumptions. My notes - and various readings - refer to famous papers that do this and the general consensus is that an exogenous 'tightening' shock in monetary policy (unexpectedly high policy rate for reasons solely stemming from the policymaker) is typically thought to cause declines in activity for some time - after some lag - along with a pickup in inflation. See the papers and IRF diagrams for more effects. These stylized facts are things you should be aware of. You should also be aware of the limitations of SVAR (structural VAR) analysis - such as requiring many identification assumptions - some of which may not be plausible, if you are trying to identify multiple shocks. It is also important that you can plausibly argue that the information set defined by the variables included in the VAR are sufficient. The 'price puzzle' mentioned in class is often though to derive from the fact that lags of s_t do not capture all the information the Fed sees. If the Fed sees, say, an increase in oil prices or gets information about oil prices not spanned by the variables already in s_t and its lags, the Fed may tighten

---

[4]Just write what $B'\Omega B$ is under our assumptions and see that the terms in the bottom row of the resulting matrix can be divided to give ϕ.

policy and yet not by enough to stop a little bit of inflation passing through for some time. Estimating a VAR inappropriately might then suggest that raising rates 'causes' a small initial positive inflation bump when that isn't actually the case. You should be aware of various other problems - especially the ones I list towards the end of the lecture.