Rational expectations

School of Economics, University College Dublin

Spring 2017

Moving beyond VAR

A VAR MODEL IS MAINLY AN ECONOMETRIC MODEL. It can help measure the shocks that hit the macroeconomy and their dynamic effects but it doesn't offer a theoretical explanation for these patterns. This requires a model with explicit dynamic and stochastic shocks. Understanding why things happen is crucial for the forecasts or analysis of policy effects.

VARs are dynamic stochastic models, but they are econometric not theoretical models and they have their limitations.

Note however that there is a very important difference in using theory as an analytical narrative or as a detailed plausible story of why things happen.

THE GOAL OF THE MODERN DYNAMIC STOCHASTIC GENERAL EQUILIBRIUM (DSGE) approach is to develop models that can explain macroeconomic dynamics as well as the VAR approach, but that are based upon the fundamental idea of optimising firms and households.

Rational expectations

A key difference between DSGE and VAR modeling is that DSGE has both backward and forward looking dynamics.

- Backward looking dynamics stem from, for instance from identities linking today's capital stock with last period's capital stock and this period's investment: $K_t = (1-\gamma)K_{t-1} + I_t$
- Forward-looking dynamics stem from optimising behaviour: What agents expect to happen tomorrow is very important for what they decide to do today.

Modeling this idea requires an assumption about how people formulate expectations. The DSGE approach relies on the idea that people have so-called *rational expectations*.

ALMOST ALL ECONOMIC TRANSACTIONS RELY on the fact that the economy is not a one-period game. Economic decisions have an intertemporal element to them. A key issue in macroeconomics is how people formulate expectations about the future in the presence of uncertainty. And so the concept of rational expectations was developed, which means two things:

Prior to the 1970s, this aspect of macro theory was largely ad hoc. Generally, it was assumed that agents used some simple extrapolative rule whereby the expected future value of a variable was close to some weighted average of its recent past values.

1. Agents use publicly available information in an efficient manner (i.e. they do not make systematic mistakes when formulating expectations).

2. Agents understand the structure of the model economy and base their expectations of variables on this knowledge.

RATIONAL EXPECTATIONS IS CLEARLY A STRONG ASSUMPTION. The structure of the economy is complex and in truth nobody truly knows how everything works. One reason for using rational expectations as a baseline assumption is that once one has specified a particular model of the economy, any other assumption about expectations means that people are making systematic errors, which seems inconsistent with rationality. But rational expectations requires one to be explicit about the full limitations of people's knowledge and exactly what kind of mistakes they make. And while rational expectations is a clear baseline, once one moves away from it there are lots of essentially ad hoc potential alternatives.

Otherwise there would also be little need for economists.

LIKE ALL MODELS, rational expectations models need to be assessed on the basis of their ability to fit the data.

First-order stochastic difference equations

Consider a very general economic model of the form

$$y_t = x_t + aE_t y_{t+1}$$

Today's value of y is determined by x and tomorrow's expected value of y. Under rational expectations the agents in the economy understand the equation and formulate their expectation in a way that is consistent with it

$E_t E_{t+1} y_{t+2} = E_t y_{t+2}$

$$\begin{aligned} E_t y_{t+1} &= E_t x_{t+1} + aE_t E_{t+1} y_{t+2} \\ E_t y_{t+1} &= E_t x_{t+1} + aE_t y_{t+2} \end{aligned}$$

This is the Law of Iterated Expectations: It is not rational for me to expect to have a different expectation next period for y_{t+2} than the one that I have today. Substituting we get

$$y_t = x_t + aE_t x_{t+1} + a^2 E_t y_{t+2}$$

Expanding to include additional periods will give

It is assumed that

$$\lim_{N\to\infty} a^N E_t y_{t+N} = 0$$

$$y_t = x_t + aE_t x_{t+1} + a^2 E_t y_{t+2} + ... + a^{N-1} E_t x_{t+N-1} + a^N E_t y_{t+N}$$
$$y_t = \sum_{k=0}^{N-1} a^k E_t x_{t+k} + a^N E_t y_{t+N}$$
$$y_t = \sum_{k=0}^{\infty} a^k E_t x_{t+k}$$

This underlies the logic of a large amount of macroeconomics.

Example: Asset pricing

CONSIDER AN ASSET which can be bought today for price P_t and which will yield a dividend of D_t. Suppose that there is a close alternative to this asset which will yield a guaranteed rate of return r. A risk neutral investor will only hold the asset if it yield the same rate of return, i.e.

$$\frac{D_t + EtP_{t+1}}{P_t} = 1 + r$$

This can be rearranged to give an expression for P_t and provide a repeated substitution solution

This equation, which states that asset prices should equal a discounted present-value sum of expected future dividends, is usually known as the dividend-discount model.

$$P_t = \frac{D_t}{1+r} + \frac{E_t P_{t+1}}{1+r}$$
$$P_t = \sum_{k=0}^{\infty} \left(\frac{1}{1+r}\right)^{k+1} E_t D_{t+k}$$

Backward solutions

$$y_t = x_t + aE_t y_{t+1}$$

Can also be written as

$$y_t = x_t + ay_{t+1} + a\epsilon_{t+1}$$

where ϵ_{t+1} is a forecast error that cannot be predicted at time t. Moving the time subscripts back one period, re-arranging, and solving via repeated substitution the backward-looking equation becomes

$$y_t = a^{-1} y_{t-1} - a^{-1} x_{t-1} - \epsilon_t$$
$$y_t = -\sum_{k=0}^{\infty} a^{-k} \epsilon_{t-k} - \sum_{k=1}^{\infty} a^{-k} x_{t-k}$$

BOTH FORWARD AND BACKWARD SOLUTIONS are correct solutions for the first-order stochastic difference equation. Which solution we use depends on the value of a.

- If $|a| > 1$, then the weights on future x_t values will explode in forward solution. Likely that the forward solution will not converge to finite sum.
- If $|a| < 1$ then the weights on the backward solution are explosive and the forward solution is the one to focus on. Knowing the path of x_t will tell you the path of y_t.

The idea that y_t depends more on values of x_t far in the distant future than on today's values is not something we can be comfortable with. In this case, practical applications should focus on the backwards solution.

The equation holds for any set of shocks ϵ_t such that $E_{t-1}\epsilon_t = 0$. The solution indeterminate. We can't actually predict what will happen with y_t even if we know the full path of x_t.

From structural to reduced form relationships

$$y_t = \sum_{k=0}^{\infty} a^k E_t x_{t+k}$$

Provides some useful insights into how y_t is determined. However, without some assumptions about how x_t evolves over time, it cannot be used to give precise predictions about the dynamics of y_t. Ideally we want to simulate the behaviour of y_t.

One reason there is a strong link between DSGE modeling and VAR models is that the question on the behaviour of y_t is often addressed by assuming that the exogenous variables x_t are generated by backward-looking time series models, like VARs. Consider the following data generating process for x_t

$$x_t = \rho x_{t-1} + \epsilon_t, \; where \; |\rho| < 1$$

This means that we have

$$E_t x_{t+k} = \rho^k x_t$$

The model's solution can be written as

$$y_t = \left[\sum_{k=0}^{\infty} (a\rho)^k \right] x_t$$

Assuming $|a\rho| < 1$, the infinite sum converges to

$$\sum_{k=0}^{\infty} (a\rho)^k = \frac{1}{1 - a\rho}$$

This identity is from the famous Keynesian multiplier formula

The model solution becomes

$$y_t = \frac{1}{1 - a\rho} x_t$$

This is the reduced-form solution for the model which, together with the equation describing the evolution of x_t, can be easily simulated on a computer.

The DSGE recipe

THE PREVIOUS SECTIONS provided a relatively simple example but illustrates the general principles for getting predictions from DSGE models

1. Obtain structural equations involving expectations of future driving variables (the $E_t x_{t+k}$ terms)

2. Make assumptions (!) about the time series process for the driving variables (x_t)

3. Solve for a reduced-form solution that can be simulated on a computer

The reduced-form model has a VAR-like representation

$$y_t = \frac{1}{1-a\rho}(\rho x_{t-1} + \epsilon_t)$$
$$= \rho y_{t-1} + \frac{1}{1-a\rho}\epsilon_t$$

Both the x_t and yt series have purely backward-looking representations. This simple model helps to explain how theoretical models tend to predict that the data can be described well using a VAR.

Example: Permanent Income Hypothesis

CONSIDER A SIMPLE PERMANENT INCOME MODEL in which consumption depends on a present discounted value of after-tax incom

$$c_t = \gamma \sum_{k=0}^{\infty} \beta^k E_t y_{t+k}$$

Suppose that the income generating process is $y_t = (1+g)Y_{t-1} + \epsilon_t$, we get the following reduced-form representation

We have $E_t y_{t+k} = (1+g)^k y_t$

$$c_t = \gamma \left[\sum_{k=0}^{\infty} (\beta(1+g))^k \right] y_t$$

Assuming that $\beta(1+g)^k < 1$, this becomes

$$c_t = \frac{\gamma}{1-\beta(1+g)} y_t$$

CONSIDERING THE MODELS IN THIS EXAMPLE; the structural equation is always true for the model; however, the reduced-form representation depends on the process for y_t taking a particular form. If that process changes, so will the reduced-form process. Lucas (1976) pointed out that the assumption of rational expectations implied that the coefficients in reduced-form models would change if expectations about the future changed and stressed that this could make reduced-form econometric models based on historical data useless for policy analysis. This problem is now known as the Lucas critique of econometric models.

Example: Temporary tax cuts

Suppose that the government is thinking about a temporary one-period income tax cut. They ask their advisers what the estimated effect of the tax cut on consumption would be. So they run a regression of consumption c_t on after-tax income y_t.

- If, in the past consumers had generally expected income growth g, then the income coefficient will be approximately $\frac{\gamma}{1-\beta(1+g)}$: each 1€increase in income produced by the tax cut will increase consumption by €$\frac{\gamma}{1-\beta(1+g)}$
- If the households have rational expectations the increase will be just €γ

The limitations of VAR analaysis

Today's DSGE models feature policy equations that describe how monetary policy is set via rules relating interest rates to inflation and unemployment; how fiscal variables depends on other macro variables; what the exchange rate regime is. These models all feature rational expectations, so changes to these policy rules will be expected to alter the reduced-form VAR-like structures associated with these economies.

This is an important selling point for modern DSGE models. These models can explain why VARs fit the data well, but they can be considered superior tools for policy analysis. They explain how reduced-form VAR-like equations are generated by the processes underlying policy and other driving variables. However, while VAR models do not allow reduced-form correlations change over time, a fully specified DSGE model can explain such patterns as the result of structural changes in policy rules.

Second-order stochastic difference equations

Variables with the following characteristic are know as jump variables

$$y_t = \sum_{k=0}^{\infty} a^k E_t x_{t+k}$$

- They only depend on what happens today and what's expected to happen tomorrow

- If expectations about the future change, they will jump
- Nothing that happened in the past restricts their movement

This may be an ok characterization for stock prices, but less so for variables in the real economy such as employment, consumption or investment. Many models in macroeconomics feature variables which depend on both the expected future values and their past values. They are characterized by second-order difference equations of the form

$$y_t = ay_{t-1} + bE_t y_{t+1} + x_t$$

Let's examine one way to solve second-order SDEs. Suppose that there is a value λ such that $v_t = y_t - \lambda y_{t-1}$ followed a first-order SDE of the form $v_t = \alpha E_t v_{t+1} = \beta x_t$. We know how to solve for v_t and back out the values for y_t. From the fact that $v_t = y_t - \lambda y_{t-1}$, we can re-write the original equation

$$\begin{aligned} v_t + \lambda y_{t-1} &= ay_{t-1} + b(E_t v_{t+1} + \lambda y_t) + x_t \\ &= ay_{t-1} bE_t v_{t+1} + b\lambda(v_t + \lambda y_{t-1} + x_t \\ (1 - b\lambda)v_t &= bE_t v_{t+1} + x_t + (b\lambda^2 - \lambda + a)y_{t-1} \end{aligned}$$

By definition, λ is a number such that the v_t it defined followed a first-order SDE. This means that λ satisfies $b\lambda^2 - \lambda + a = 0$. Given that this is a quadratic equation, there are two values for λ that satisfy it, and for each of these values we can characterize v_t by

Usually only one of the potential values of λ is less than one in absolute value, so this gives the unique stable solution.

$$\begin{aligned} v_t &= \frac{b}{1 - b\lambda} E_t v_{t+1} + \frac{1}{1 - b\lambda} x_t \\ &= \frac{1}{1 - b\lambda} \sum_{k=0}^{\infty} \left(\frac{b}{1 - b\lambda} \right)^k E_t x_{t+k} \end{aligned}$$

And y_t obeys

$$y_t = \lambda y_{t-1} + \frac{1}{1 - b\lambda} \sum_{k=0}^{\infty} \left(\frac{b}{1 - b\lambda} \right)^k E_t x_{t+k}$$

THE LOGIC UNDERLYING the first-order SDE underlies the solution methodology for just about all rational expectations models. Suppose that we have a vector of variables

$$Z_t = \begin{pmatrix} z_{1t} \\ z_{2t} \\ . \\ z_{nt} \end{pmatrix}$$

This can be represented in a macroeconometric model of the form

B is an nxn matrix

$$Z_t = BE_t Z_{t+1} + X_t$$

The logic of repeated substitution can be applied to this model to give a solution of the form

$$Z_t = \sum_{k=0}^{\infty} B^k E_t X_{t+k}$$

As with the single equation model, this will only give a stable solution under certain conditions. And for this we need to consider eigenvalues.

λ_i is an eigenvalue of the matrix B if there exists a vector e_i such that $Be_i = \lambda_i e_i$. Many nxn matrices have n distinct eigenvalues. Denote by P the matrix that has as columns n eigenvectors corresponding to these eigenvalues. In this case $BP = P\Omega$, Ω is a diagonal matrix of eigenvalues.

$$\Omega = \begin{pmatrix} \lambda_1 & 0 & 0 & 0 \\ 0 & \lambda_2 & 0 & 0 \\ 0 & 0 & . & 0 \\ 0 & 0 & 0 & \lambda_n \end{pmatrix}$$

This implies that $B = P\Omega P^{-1}$, which tells us something about the relationship between eigenvalues and higher powers of B because

$$B^n = P\Omega P^{-1} = P \begin{pmatrix} \lambda_1 & 0 & 0 & 0 \\ 0 & \lambda_2 & 0 & 0 \\ 0 & 0 & . & 0 \\ 0 & 0 & 0 & \lambda_n \end{pmatrix} P_{-1}$$

- The difference between lower and higher powers of B is that the higher powers depend on the eigenvalues take to the power n
- If all of the eigenvalues are in the unit circle, i.e. less than one in absolute value, then all of the entries in B^n will tend towards zero as $n \to \infty$

So a condition that ensures that model $Z_t = BE_t Z_{t+1} + X_t$ has a unique stable forward-looking solution is that the eigenvalues of B are all inside the unit circle.

How to calculate eigenvalues

Consider the following 2 x 2 matrix

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

Suppose that A has two eigenvalues, λ_1 and λ_2, define vector λ as

$$\lambda = \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix}$$

The fact that there are eigenvectors when multiplied by $A - \lambda I$ equal a vector of zeroes means that the determinant of the matrix equal zero.

$$A - \lambda I = \begin{pmatrix} a_{11} - \lambda & a_{12} \\ a_{21} & a_{22} - \lambda_2 \end{pmatrix}$$

Solving the quadratic formula will give the two eigenvalues of A

$$(a_{11} - \lambda_1 a_{12})(a_{22} - \lambda_2) - a_{12}a_{21} = 0$$

The Binder-Pesaran model

Consider the following vector Z_t

$$Z_t = AZ_{t-1} + BE_t Z_{t+1} + HX_t$$

The restriction to one-lag one-lead form is only apparent, and the companion matrix trick can be used to allow this model to represent models with n leads and lags. In this sense, this equation summarizes all possible linear rational expectations models.

Binder and Pesaran solved this model in a manner exactly analogous to the second-order SDE: Find matrix C such that $W_t = Z_t - CZ_{t-1}$ obeys a first-order matrix of the form $W_t = FE_t W_{t+1} + GX_t$.

We transform the problem of solving the second-order system in equation into a simpler first-order system.

What must C be? Using the fact that $Z_t = W_t + CZ_{t-1}$ we can re-write the model as

$$\begin{aligned} W_t + CZ_{t-1} &= AZ_{t-1} + B(E_t W_{t+1} + CZ_t) + HX_t \\ &= AZ_{t-1} + B(E_t W_{t+1} + C(W_t + CZ_{t-1})) + HX_t \end{aligned}$$

This re-arranges to

$$(I - BC)W_t = BE_t W_{t+1} + (BC^2 - C + A)Z_{t-1} + HX_t$$

Given that C is a matrix that makes W_t follow a first-order forward-looking matrix equation, i.e. no extra Z_{t-1} terms, we get

$$BC^2 - C + A = 0$$

This equation can be solved to give C One methods to do this is using the fact that $C + BC^2 + A$

1. Provide initial guess $C_0 = I$

2. Iterate on $C_n = BC_{n-1}^2$ untill al entries in C_n converge

Once we know C we have

$$W_F E_t W_{t+1} + GX_t$$

where

$$F = (I - BC)^{-1} B$$
$$G = (I - BC)^{-1} H$$

Assuming that all eigenvalues of F are inside the unit circle, this has a stable forward-looking solution

$$W_t = \sum_{k=0}^{\infty} F^k E_t(GX_{t+k})$$

Which can be written in terms of the original equation as

$$Z_t = CZ_{t-1} + \sum_{k=0}^{\infty} F^k E_t(GX_{t+k})$$

Reduced-form representation

Suppose that the driving varialbes X_t follow a VAR representation of the form

$$X_D X_{t-1} + \epsilon_t$$

where the eigenvalues for D are inside the unit circle. This implies $E_t X_{t+l} = D^k X_t$, so the model solution is

$$Z_t = CZ_{t-1} + \left[\sum_{k=0}^{\infty} F^k G D^k \right] X_t$$

The infinite sum in this equation will converge to matrix P to the model has a reduced form representation which can be simulated along with the VAR process for the driving variables.

$$Z_t = CZ_{t-1} + PX_t$$

This provides a relatively simple recipe for simulating DSGE models:

1. Specify matrices A, B, H
2. Solve for C, F, G
3. Specify a VAR process for the driving variables
4. Obtain the reduced-form representations

The equations we get from model will often contain multiple values of different variables at time t. This is not a problem as we can plug the model into a computer as

$$KZ_t = AZ_{t-1} + BE_tZ_{t+1} + HX_t$$

The program can multiply both sides by K^{-1} and solved using the Binder-Pesaran model.

$$Z_t = K^{-1}AZ_{t-1} + K^{-1}BE_tZ_{t+1} + K^{-1}HX_t$$

This might seem a bit complicated, but in practice it is not so hard. You figure out what your model implies in terms of the matrices K, A, B, H and the computer will give you the following representation which can be used for calculations.

$$\begin{aligned} Z_t &= CZ_{t-1} + PX_t \\ X_t &= DX_{t-1} + \epsilon_t \end{aligned}$$