Estimating DSGE Models

School of Economics, University College Dublin

Spring 2017

To understand the technique with which modern DSGE models are estimated, a number of issues need to be covered, including

1. Role of shocks in the model
2. Observable and unobservable variables
3. Kalman filter estimation of state-space models
4. Bayesian methods

Starting point: a solved model

The estimation starts with the solved version of a log-linearised model. Suppose that we have a model described by

Z_t is a set of n endogenous variables and X_t is a set of k exogenous variables that evolve according to $X_t = DX_{t-1} + \epsilon_t$.

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

This model has a solution of the form

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

C depends on the coefficients in A, B and P depends on the coefficients in A, B, H, D.

To establish the properties of the model we could simulate it. However, we are interested in the estimates of the coefficients. In particular, using the observed data we want to know the estimates of the coefficients in A, B, D, H. The estimation here depends on the model and the kind of data that we have. Consider the case where X_t, Z_t are observable. Here the model makes a clear prediction that given any set of structural parameter A, B, H, D, the data will be given by

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

However, is is likely that there actually does not exist a set of A, B, D, H matrices that will allow the model to fit the data perfectly. This is due to cross-equation restrictions in DSGE models which are very limiting: given the matrices particular patterns must be obeyed by the C, P matrices. Since the model will not fit the data, we can't use Maximum Likelihood Estimation (MLE). This problem could be addressed by adding error terms u_t. A, B, D, H can be estimated in that case, and the best fitting model will have a form such as

Recall that MLE is a method that returns estimates on the basis of how likely the model might fit the data. In this case we know for sure that the model does not fit the data.

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

u_t does not have any microeconomic foundation, but it will provide a sense of how well the model fits the data.

Maximum Likelihood Estimation

Let's consider estimating a model with observable variables using MLE.

$$Z_t = CZ_{t-1} + PX_t + u_t$$
$$X_t = DX_{t-1} + \epsilon_t$$

$u_t \sim N(0, \sum_u)$, $\epsilon_t \sim N(0, \sum_\epsilon)$

Suppose that the endogenous variables are observed by $Z_1, Z_2, ..., Z_T$ and the exogenous variables by $X_1, X_2, ..., X_T$. Here we can combine the log-likelihood functions for the Z and X data as the likelihood of the full model multiplies the likelihood of the X data and the likelihood of the Z data. Therefore, the maximum likelihood estimates of $A, B, H, D, \sum_\epsilon, \sum_u$ are those that maximise the following log-likelihood

$$-T\ log 2\pi - T(log|\sum_\epsilon^{-1}| + log|\sum_u^{-1}|)$$
$$-\frac{1}{2}\sum_{k=1}^{T}(X_i - DX_{i-1})'\sum_\epsilon^{-1}(X_i - DX_{i-1})$$
$$-\frac{1}{2}\sum_{k=1}^{T}(Z_i - CZ_i - PX_i)'\sum_u^{-1}(Z_i - CZ_i - PX_i)$$

Subject to the restrictions that map A and B into C and map A, B, H, D into P.

A mix of observables and unobservables

A complication of most DSGE model is that they do not exclusively rely on observable variables, but often they are a mix of observable and unobservable variables. To illustrate this, let's consider the Real Business Cycle (RBC) model that we have discussed before. This model was describer by the following equations.

$$y_t = \left(1 - \frac{\alpha\gamma}{\beta^{-1} + \gamma - 1}\right) c_t + \left(\frac{\alpha\gamma}{\beta^{-1} + \gamma - 1}\right) i_t$$
$$y_t = a_t + \alpha k_{t-1} + (1-\alpha) n_t$$
$$k_t = \gamma i_t + (1-\gamma) k_{t-1}$$
$$n_t = y_t - \eta c_t$$
$$c_t = E_t c_{t+1} - \frac{1}{\eta} E_t r_{t+1}$$
$$r_t = (1 - \beta(1-\gamma))(y_t - k_{t-1})$$
$$a_t = \rho a_{t-1} + \epsilon_t$$

This model features 7 equations with 6 endogenous variables and one exogenous variable

- Endogenous: $y_t, c_t, i_t, k_t, n_t, r_t$

- Exogenous: a_t

The model also mixes 4 observable variables and three unobservable variables

- Observable: y_t, c_t, i_t, n_t
- Unobservable: a_t, k_t, n_t

In order to deal with the unobservable variables, special techniques are required.

Stochastic singularity problem

Models such as the RBC model provide a micro-foundation for why the observed data can be perfectly fitted to a model. The main issue is that there is an unobservable technology series upon which all the observed series depend. While the model features stochastic shocks, it also has a feature that is know as stochastic singularity. This means that the shocks in all the equations are just multiples of each other. The model will therefore predict that certain ratios of the observed variables will be constant, while in reality these predictions do not hold. Meaning that the model will not fit the data. For a model to have well-defined econometric estimates it is therefore necessary that for every observable variable there is at least one unobservable shock. This can take in shape in two different forms

This type of model cannot be estimated with MLE since the unobserved series shows up in all the reduced-form solution equations.

e.g. current and lagged consumption of investment.

1. measurement error
2. involve a shock in each equation with a clear structural interpretation

DSGEs are state-space models

A DSGE model that is a mix of observable and unobservable variables is an example of state-space model. This type model can be described using two equations

1. the state (transition) equation S_t
2. the measurement equation Z_t

$$S_t = FS_{t-1} + u_t$$
$$Z_t = HS_t + v_t$$

The state equation describes how a set of unobservable state variables evolves over time. The measurement equation related the observable variables to the unobservable state variables.

The error terms u_t, v_t can include normally distributed errors or zeroes if the equation described is an identity.

State-space models

State-space models are a general class of linear time series models that combine observable and unobservable variables. And as discussed in the previous section they can be described by two equations for state and measurement. In a state-space model the observed data is thus described by

$$Z_t = HS_t + v_t$$

Now S_t can't be observed, but we can replace it with an observable unbiased guess based on the information available up to time $t-1$: $S_t|t-1$. We could therefore rewrite the measurement equation as

$$Z_t = HS_{t|t-1} + v_t + H(S - S_{t|t-1})$$

Assume that the errors are normally distributed with a know covariance matrix

$$S_t - S_{t-1} \sim N(0, \textstyle\sum_{t|t-1}^{S})$$

$S_{t|t-1}$ is observable and since the unobservable elements v_t and $S_t - S_{t|t-1}$ are normally distributed, this model can be estimate using maximum likelihood estimation.

The Kalman filter

MLE can be used to estimate a model where we include an unbiased guess for S_t. Now all we need is a method to generate the unbiased guess, and one approach that is commonly used for this is the Kalman filter. The Kalman filter is an iterative method where one starts with the estimate of the state variables in one period and uses observable data from the next period to update these variables. To estimate the state variable at any given time t given information at time $t-1$ we use

Not entirely unlike Bayesian methods.

$$\begin{aligned} S_t &= FS_{t-1} + u_t \\ S_{t|t-1} &= FS_{t-1|t-1} \end{aligned}$$

This means that in period $t-1$ the expected value for the observable variables at time t are

$$Z_{t|t-1} = HS_{t|t-1} = HFS_{t-1|t-1}$$

The next step here is to update the state variable in time t when we observe Z_t given the information we get from $Z_{t|t-1} = HFS_{t-1|t-1}$. Here we will rely on conditional expectations. The assumptions of the model implies the following

$$\begin{pmatrix} S_t \\ Z_t \end{pmatrix} \sim N\left(\begin{pmatrix} FS_{t-1|t-1} \\ HFS_{t-1|t-1} \end{pmatrix}, \begin{pmatrix} \sum_{t|t-1}^{S} & \left(H\sum_{t|t-1}^{S}\right)' \\ \left(H\sum_{t|t-1}^{S}\right) & \sum^{V} + H\sum_{t|t-1}^{S} H' \end{pmatrix}\right)$$

At this stage we can introduce conditional expectations to state that the minimum variance unbiased estimate of S_t given observed Z_t will be

$$E(S_t|Z_t) = S_{t|t} = FS_{t-1|t-1} + K_t(Z_t - HFS_{t-1|t-1})$$

$$K_t = \left(H\sum\nolimits^{S}_{t|t-1}\right)'\left(\sum\nolimits^{V} + H\sum\nolimits^{S}_{t|t-1} H'\right)^{-1}$$

The covariance matrices to compute K_t are updated by the formulae

$$\sum\nolimits^{S}_{t|t-1} = F\sum\nolimits^{S}_{t-1|t-1} F' + \sum\nolimits^{U}$$
$$\sum\nolimits^{S}_{t|t} = (I - K_t H)\sum\nolimits^{S}_{t|t-1}$$

All we need now is an initial estimate for $S_{1|0}$ as well as the covariance matrix to start the filtering process. For most macroeconomic models the state variables can be assumed to have zero mean, which serves as the initial guess. To estimate the unconditional variance we can use the estimate of the variance of an estimate of the state variable from a large data sample. Recall that

$$\sum\nolimits^{S}_{t|t-1} = F\sum\nolimits^{S}_{t-1|t-1} F' + \sum\nolimits^{u}$$

The values of the covariance matrix generated by this equation will generally converge, this means that for the unconditional covariance matrix a $\sum$ value can be used which solves

$$\sum = F\sum F' + \sum^{u}$$

The Kalman filter is a one-sides filter as the estimate of the state variable at time t are based only on information available at time t. So no data after period t is used to calculate the estimates of the unobserved state variables. This is a reasonable model when you are studying the state variable in real time. However, for most studies data is available for after time t which can be used to estimate time-varying models using a Kalman smoother. This is a two-sided filter that uses data both before and after t to compute the expected values of the state variables at time t.

Example: an RBC model

Going back to the example of the RBC model, the solution to the RBC model can, excluding labour, be summarised as

$$\begin{aligned} k_t &= a_{kk}k_{t-1} + a_{kz}z_t \\ c_t &= a_{ck}k_{t-1} + a_{cz}z_t \\ z_t &= \rho z_{t-1} + \epsilon_t \end{aligned}$$

For the sake of the illustration let's assume that consumption and capital are only observed with error, this makes that the two observable variables are

$$\begin{aligned} k^*_t &= a_{kk}k_{t-1} + akzz_t + u^k_t \\ c^*_t &= a_{ck}k_{t-1} + aczz_t + u^c_t \end{aligned}$$

This can be written in state-space form using the transition and measurement equation

$$\begin{pmatrix} k_{t-1} \\ z_t \end{pmatrix} = \begin{pmatrix} a_{kk} & a_{kz} \\ 0 & \rho \end{pmatrix} \begin{pmatrix} k_{t-2} \\ z_{t-1} \end{pmatrix} + \begin{pmatrix} 0 \\ \epsilon_t \end{pmatrix}$$
$$\begin{pmatrix} k^*_{t-1} \\ c^*_t \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ a_{ck} & a_{cz} \end{pmatrix} \begin{pmatrix} k_{t-1} \\ z_t \end{pmatrix} + \begin{pmatrix} u^k_{t-1} \\ u^c_t \end{pmatrix}$$

Here we get that

Some adjustments had to be made to get the model in state-space form and the timing conventions associated with this representation are not quite the same as in the original model.

$$S_t = \begin{pmatrix} k_{t-1} \\ z_t \end{pmatrix}$$
$$Z_t = \begin{pmatrix} k^*_{t-1} \\ c^*_t \end{pmatrix}$$

This can be used for all DSGE models which can be estimated using the Kalman filter.

Estimating DSGE with MLE (again)

Using the Kalman filter one can actually estimate a DSGE model that mixes observable and unobservable variables using MLE. Once you have specified the model you can feed it into a computer package which will

1. Sort the model into space-state methods
2. Find possible parameter values
3. Use Kalman filter to smooth parameters
4. Produce period-by-period likelihoods for possible parameter values
5. Pick best parameters and calculate standard errors (using MLE)

Using MLE is not without dangers however. Some of the issues with maximising the likelihood of a DSGE include

- Large number of parameters
- Sparsity of data (often quarterly)
- Flexible nature of DSGE, generating similar behaviour with different parameter values
- Standard errors difficult to compute

Due to this difficulties, most current research prefers to use Bayesian methods. The Bayesian approach specifies a prior likelihood which will be combined with the likelihood function to produce an estimate of the posterior. Once the posterior is calculated it is relatively straightforward to produce means, confidence intervals, etc. The main advantage of the Bayesian approach over MLE is that it uses a full likelihood function rather than a single point estimate. One important question concerning the use of Bayesian method is of course how the prior is determined. In practice the prior is defined using a distribution in the form that fits common sense and corresponds to previous studies.

Example: the Smets-Wouters model

The supply side

The aggregate production function is given by

$$y_t = \phi_p(\alpha k_t^s + (1-\alpha)l_t + \epsilon_t^a)$$

y_t is GDP, k_t^s is capital in use ,l_t is labour input, ϵ_t^a is total factor productivity

In this model capital in use k_t is determined by the lagged level of capital and a capacity utilisation variable

$$k_t^s = k_{t-1} + z_t$$

This capacity utilisation variable is linked to the marginal productivity of capital since there are costs associated with adjusting the amount of capital in use. The marginal productivity of capital itself is a function of the capital to labour ratio and the real wage

$$r_t^k = -(k_t - l_t) + w_t$$

Total factor productivity will evolve over time according to

$$\epsilon_t^a = \rho\epsilon_{t-1}^a + \eta_t^a$$

The demand side

The model includes the following resource constraint

$$y_t = c_y c_t + i_y i_t + z_y z_t + \epsilon_t^g$$

y_t is GDP, c_t is consumption, i_t is investment, and ϵ_t^g is exogenous spending. Variables with subscript y are constants.

z_t is included in the resource constraint because of the assumption that there are costs associated with having high rates of capital utilisation. Exogenous spending is assumed to develop over time according to

$$\epsilon_t^g = \rho\epsilon_{t-1}^g + \eta_t^g + \rho_{ga}\eta_t^a$$

Exogenous spending is assumed to have two components, i) government spending, ii) an element related to productivity.

Consumption

Consumption is determined by

$$c_t = c_1 c_{t-1} + (1 - c_1) E_t c_{t+1} + c_2(l_t - E_t l_{t+1}) - c_3(r_t - E_t \pi_{t+1} + \epsilon_t^b)$$

c_1, c_2, c_3 are constant parameters (themselves functions of deeper structural parameters), r_t is the interest rate on a one-period safe bond (quarterly).

Here ϵ_t^b develops over time according to

$$e_t^b = \rho_b \epsilon_{t-1} + \eta_t^b$$

ϵ_b is a risk premium shock determining the willingness of a household to hold the one-period bond. This can also be seen as a type of preference shock that influence short-term consumption-saving decisions.

Some other things to be aware of with regard to the equation for consumption include

- The backward looking consumption term represent habit forming
- The equation allows for substitution of consumption with labour input

Investment

Investment is determined by

$$i_t = i_t i_{t-1} + (1 - i_1) E_t i_{t+1} + i_2 q_t + \epsilon_t^i$$

Similar to consumption we can see that investment depends on its lagged value. In this case because there is an adjustment cost function that limits the amount of new investment that are immediately available. Th investment level is mainly driven by q_t which is described by

$$q_t = q_1 E_t q_{t+1} + (1 - q_1) r_{t+1}^k - (r_t - E_t \pi_{t+1} + \epsilon_t^b)$$

$k_t = k_1 k_{t+1} + (1 - k_1) i_t + k_2 \epsilon_t^i$

Here q_t depends positively on the expected future marginal productivity of capital and negatively on the future real interest rate.

Prices

The mark up of price over marginal cost is determined by

$$\mu_t^p = \alpha(k_t - l_t) + \epsilon_t^a - w_t$$

This equation accounts for the diminishing marginal productivity of capital, the effects of a productivity shock on costs and the real wage. Price inflation is given by

$$\pi_t = \pi_1 \pi_{t-1} + \pi_2 E_t \pi_{t+1} - \pi_3 \mu_t^p + \epsilon_t^p$$

As described in the paper, this is a New-Keynesian Phillips curve. However, it is adjusted to account for lagged inflation. This is done based on the assumption that most firms will index their prices based on past inflation levels and can only set an optimal price occasionally, following the Calvo model. ϵ_t^p is a price mark-up disturbance which is described by

$$\epsilon_t^p = \rho^p \epsilon_{t-1}^p + \eta_t^p - \mu_p \eta_{t-1}^p$$

This shock affects both current and lagged inflation in order to get a temporary price level shock.

Wages

Wages are given by

$$w_t = w_1 w_{t-1} + (1 - w_1) E_t(w_{t+1} + \pi_{t+1}) - w_2 \pi_t + w_3 \pi_{t-1} - w_t \mu_t^w + \epsilon_t^w$$

Where the shock is described by

$$\epsilon_t^w = \rho^w \epsilon_{t-1}^w + \eta_t^w - \mu_w \eta_{t-1}^w$$

As you can see, wages are largely determined by past wages and inflation but also by the μ_t^w term. This term is the wage mark-up which is the gap between the real wage and the marginal rate of substitution between working and consuming or

$$\begin{aligned}\mu_t^w &= w_t - mrs_t \\ &= w_t - \left(\sigma l_t - \frac{1}{1 - \lambda/\gamma}(c_t - \lambda c_{t-1})\right)\end{aligned}$$

Basically we have something like sticky wages here. The wages adjust gradually to equate the marginal costs and benefits of working.

Monetary policy

Concerning monetary policy it is assumed that the central banks sets the short-term interest rates according to

$$r_t = \rho r_{t-1} + (1 - \rho)(r_\pi \pi_t + r_y(y_t - y_t^p)) + r_{\Delta y}[(y_t - y_t^p) - (y_{t-1} - y_{t-1}^p)] + \epsilon_t^r$$

$$\epsilon_t^r = \rho^r \epsilon_{t-1}^r + \eta_t^r$$

The interest rate depends on last period's interest rate while gradually adjusting towards a target interest rate that depends on inflation and the gap between output and its potential level, as well as the growth rate of the output gap.

Potential output is defined as the level of output that would prevail if prices and wages were fully flexible. This means the model effectively needs to be expanded to add a shadow flexible-price economy.

Final model

The observable VAR system is given by

$$Y_t = \begin{pmatrix} dlGDP_t \\ dlCONS_t \\ dlINV_t \\ dlWG_t \\ lHOURS_t \\ dlP_t \\ FEDFUNDS_t \end{pmatrix} = \begin{pmatrix} \overline{\gamma} \\ \overline{\gamma} \\ \overline{\gamma} \\ \overline{\gamma} \\ \bar{l} \\ \overline{\pi} \\ \bar{r} \end{pmatrix} + \begin{pmatrix} y_t - y_{t-1} \\ c_t - c_{t-1} \\ i_t - it - 1 \\ w_t - w_{t-1} \\ l_t \\ \pi_t \\ r_t \end{pmatrix}$$

Compared to the standard New-Keynesian Model (or RBC model) the model includes a lot of additional features such as

- Adjustment costs for investment
- Capacity utilisation cost
- Habit persistence
- Price indexation
- Wage indexation
- All kinds of autocorrelated shock terms

This fixes are mainly included in order to overcome shortcoming of previous models and basically to slow things down to give random shocks longer lasting effects and making the development of variables more sluggish. The wage and price indexation is used in order to overcome the failure of the New Keynesian model to deal with persistence in inflation. Note that these adjustments are largely ad hoc and don't have a clear theoretical grounding.

Recall that this was a major shortcoming for the RBC model.

Results

Strengths and weaknesses of DSGE models

Summarising, some weaknesses of DSGE models are

1. Large number of ad hoc mechanisms
   - Included mainly to fit the data rather than following from theory
2. Large amount of unexplained, often highly persistent, shocks
3. Little attention for financial markets
4. Limited modeling of policy tools or details of national accounts

Figure 1: Prior and posterior distribution on the structural parameters (top) and shock processes (bottom).

TABLE 1A—PRIOR AND POSTERIOR DISTRIBUTION OF STRUCTURAL PARAMETERS

	Prior distribution			Posterior distribution			
	Distr.	Mean	St. Dev.	Mode	Mean	5 percent	95 percent
φ	Normal	4.00	1.50	5.48	5.74	3.97	7.42
σ_c	Normal	1.50	0.37	1.39	1.38	1.16	1.59
h	Beta	0.70	0.10	0.71	0.71	0.64	0.78
ξ_w	Beta	0.50	0.10	0.73	0.70	0.60	0.81
σ_l	Normal	2.00	0.75	1.92	1.83	0.91	2.78
ξ_p	Beta	0.50	0.10	0.65	0.66	0.56	0.74
ι_w	Beta	0.50	0.15	0.59	0.58	0.38	0.78
ι_p	Beta	0.50	0.15	0.22	0.24	0.10	0.38
ψ	Beta	0.50	0.15	0.54	0.54	0.36	0.72
Φ	Normal	1.25	0.12	1.61	1.60	1.48	1.73
r_π	Normal	1.50	0.25	2.03	2.04	1.74	2.33
ρ	Beta	0.75	0.10	0.81	0.81	0.77	0.85
r_y	Normal	0.12	0.05	0.08	0.08	0.05	0.12
$r_{\Delta y}$	Normal	0.12	0.05	0.22	0.22	0.18	0.27
$\bar{\pi}$	Gamma	0.62	0.10	0.81	0.78	0.61	0.96
$100(\beta^{-1} - 1)$	Gamma	0.25	0.10	0.16	0.16	0.07	0.26
$\bar{l}$	Normal	0.00	2.00	−0.1	0.53	−1.3	2.32
$\bar{\gamma}$	Normal	0.40	0.10	0.43	0.43	0.40	0.45
α	Normal	0.30	0.05	0.19	0.19	0.16	0.21

TABLE 1B—PRIOR AND POSTERIOR DISTRIBUTION OF SHOCK PROCESSES

	Prior distribution			Posterior distribution			
	Distr.	Mean	St. Dev.	Mode	Mean	95 percent	5 percent
σ_a	Invgamma	0.10	2.00	0.45	0.45	0.41	0.50
σ_b	Invgamma	0.10	2.00	0.24	0.23	0.19	0.27
σ_g	Invgamma	0.10	2.00	0.52	0.53	0.48	0.58
σ_I	Invgamma	0.10	2.00	0.45	0.45	0.37	0.53
σ_r	Invgamma	0.10	2.00	0.24	0.24	0.22	0.27
σ_p	Invgamma	0.10	2.00	0.14	0.14	0.11	0.16
σ_w	Invgamma	0.10	2.00	0.24	0.24	0.20	0.28
ρ_a	Beta	0.50	0.20	0.95	0.95	0.94	0.97
ρ_b	Beta	0.50	0.20	0.18	0.22	0.07	0.36
ρ_g	Beta	0.50	0.20	0.97	0.97	0.96	0.99
ρ_I	Beta	0.50	0.20	0.71	0.71	0.61	0.80
ρ_r	Beta	0.50	0.20	0.12	0.15	0.04	0.24
ρ_p	Beta	0.50	0.20	0.90	0.89	0.80	0.96
ρ_w	Beta	0.50	0.20	0.97	0.96	0.94	0.99
μ_p	Beta	0.50	0.20	0.74	0.69	0.54	0.85
μ_w	Beta	0.50	0.20	0.88	0.84	0.75	0.93
ρ_{ga}	Beta	0.50	0.20	0.52	0.52	0.37	0.66

Figure 2: Out-of-sample prediction performance showing that the DSGE model outperforms the VAR and BVAR models.

TABLE 3—OUT-OF-SAMPLE PREDICTION PERFORMANCE

	GDP	dP	Fedfunds	Hours	Wage	CONS	INV	Overall
VAR(1)	*RMSE-statistic for different forecast horizons*							
1q	0.60	0.25	0.10	0.46	0.64	0.60	1.62	−12.87
2q	0.94	0.27	0.18	0.78	1.02	0.95	2.96	−8.19
4q	1.64	0.34	0.36	1.45	1.67	1.54	5.67	−3.25
8q	2.40	0.53	0.64	2.13	2.88	2.27	8.91	1.47
12q	2.78	0.63	0.79	2.41	4.09	2.74	10.97	2.36
BVAR(4)	*Percentage gains (+) or losses (−) relative to VAR(1) model*							
1q	2.05	14.14	−1.37	−3.43	2.69	12.12	2.54	3.25
2q	−2.12	15.15	−16.38	−7.32	−0.29	10.07	2.42	0.17
4q	−7.21	31.42	−12.61	−8.58	−3.82	1.42	0.43	0.51
8q	−15.82	33.36	−13.26	−13.94	−8.98	−8.19	−11.58	−4.10
12q	−15.55	37.59	−13.56	−4.66	−15.87	−3.10	−23.49	−9.84
DSG	*Percentage gains (+) or losses (−) relative to VAR(1) model*							
1q	5.68	2.05	−8.24	0.68	5.99	20.16	9.22	3.06
2q	14.93	10.62	−17.22	10.34	6.20	25.85	16.79	2.82
4q	20.17	46.21	1.59	19.52	9.21	26.18	21.42	6.82
8q	22.55	68.15	28.33	22.34	15.72	21.82	25.95	11.50
12q	32.17	74.15	40.32	27.05	21.88	23.28	41.61	13.51

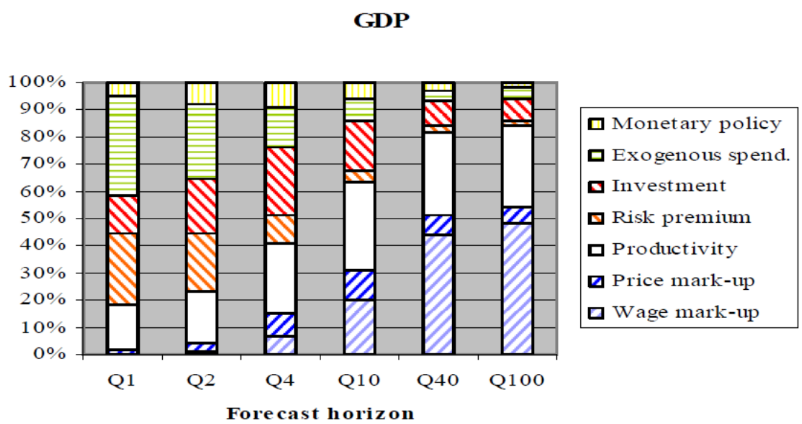
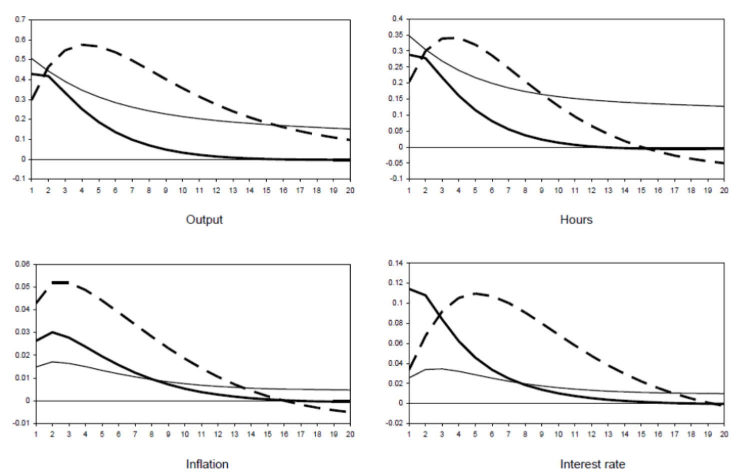


Figure 3: Forecast error variance decomposition for GDP at different forecast horizons. In short run GDP movements are driven by exogenous spending, the risk premium shock, and the investment-specific technology shock. These can be classified as demand shocks, see figure 4. In the medium and long run productivity and wage mark-up shock (supply side) account for most of the variation, see figure 5.



Notes: Bold solid line: risk premium shock; thin solid line: exogenous spending shock; dashed line: investment shock.

Figure 4: Estimated mean impulse responses to demand shocks. Bold solid line is the risk premium shock; thin solid line is exogenous spending; dashed line investment shock.

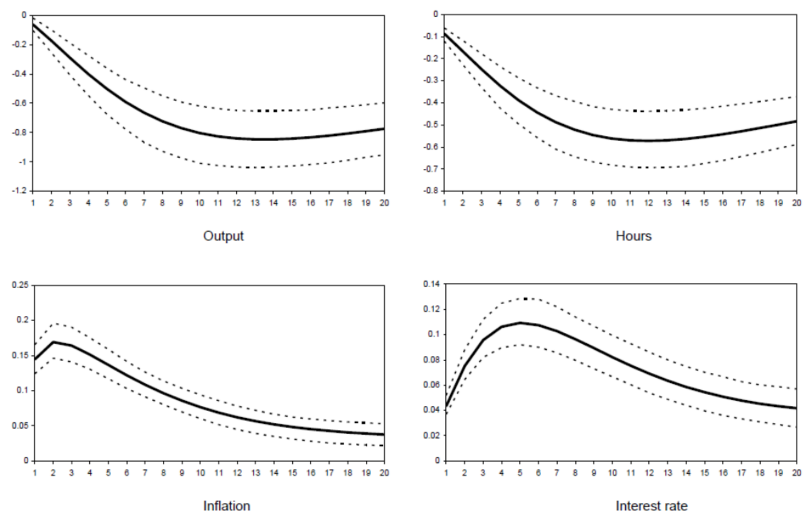


Figure 5: Estimated impulse response to a wage mark-up shock. Solid line is the mean, dotted lines are 10 and 90% posterior intervals. A positive wage mark-up shock gradually reduces output and hours worked.

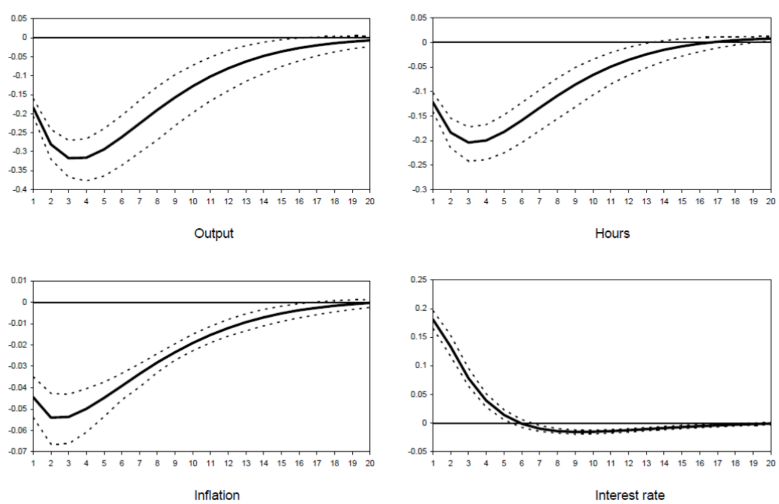


Figure 6: Impulse response to a monetary policy shock. Solid line is the mean, dotted lines are 10 and 90% posterior intervals. Peak effect of policy shock on inflation occurs before its peak effect on output.

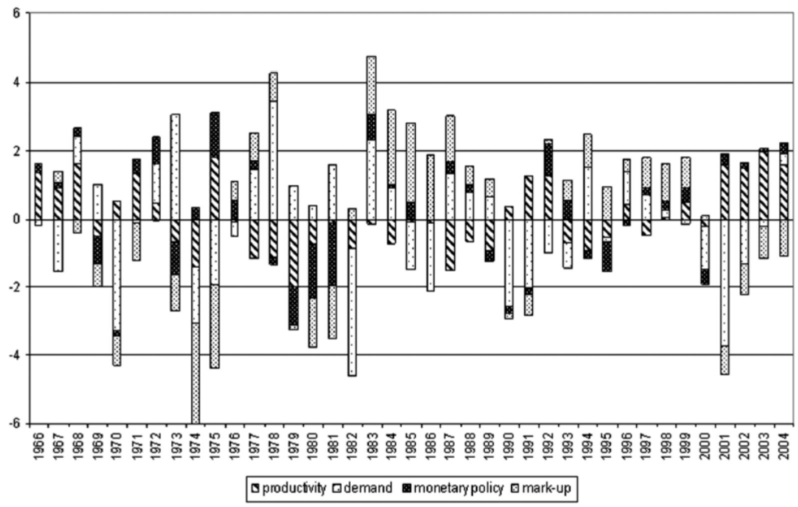


Figure 7: Historical decomposition of GDP growth.

5. Pure rational expectations might not hold up
6. Not really based on stable structural parameters (i.e. not immune to the Lucas critique)

The strengths of DSGE model vis-a-vis VAR models include

1. Imposing budget constraints
2. Theoretically grounded argument for how agents behave
3. Coherent handling of expectations
4. More suitable for forecasting and what-if analyses

Background: Conditional expectations

Suppose that we are interested in getting the estimate of the value of variable X, but the problem is that we don't actually observe this variable. We can observe variable Z which is correlated with X. Specifically, let's assume that X and Z are jointly normally distributed

$$\begin{pmatrix} X \\ Z \end{pmatrix} \sim N\left(\begin{pmatrix} \mu_X \\ \mu_Z \end{pmatrix}, \begin{pmatrix} \sigma_X^2 & \sigma_{XZ} \\ \sigma_{XZ} & \sigma_X^2 \end{pmatrix}\right)$$

Since we can observe Z which is correlated with X the expected value of X conditional on observing Z becomes

$$E(X|Z) = \mu_X + \frac{\sigma_{XZ}}{\sigma_Z^2}(Z - \mu_Z)$$

The weight you put on the information in Z when formulating an expectation of X depends on two things

1. The level of correlation with Z

2. The relative standard deviation

If Z has a high standard deviation then you don't place much weight on it. We can extend this from two to more variables. In this case X will be a $1xn$ vector and Z a $1xm$ vector. If we assume that all variables are jointly normally distributed we get

$$\begin{pmatrix} X \\ Z \end{pmatrix} \sim N\left(\begin{pmatrix} \mu_X \\ \mu_Z \end{pmatrix}, \begin{pmatrix} \sum_{XX} & \sum_{XZ} \\ \sum_{XZ} & \sum_{ZZ} \end{pmatrix}\right)$$

The expected value of X conditional on observing Z becomes

$$E(X|Z) = \mu_X + \sum_{XZ} \sum_{ZZ}^{-1}(Z - \mu_Z)$$