# Estimating DSGE models

School of Economics, University College Dublin

Spring 2018

DSGE models are complex; early models were calibrated

- Picking parameter values that match steady-state values
  - Labour share of income, capital-output ratio, etc.
- Combine with historical averages
- Parameter estimates from microeconomic studies
  - Relative risk aversion, labour supply elasticities, depreciation rates

Alternatively there was **indirect inference**: choosing parameters to match certain data moments

Number of issues to cover concerning DSGE model estimation

1. Divide model into observable and unobservable variables
2. Role played by shocks in the model
3. State-space models estimated with Kalman filter
4. Bayesian methods

## Solved model

$$KZ_t = AZ_{t-1} + B\mathbb{E}_t Z_{t+1} + HX_t \quad (1)$$

$Z_t$ is a set of $n$ endogenous variables
$X_t$ is a set of $k$ exogenous variables; evolves according to

$$X_t = DX_{t-1} + \epsilon_t \quad (2)$$

Solution of model given by

$$Z_t = CZ_{t-1} + PX_t \quad (3)$$

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

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

To establish model's properties could simulate it

- However, of interest are coefficient estimates in $A, B, D, H$
- Estimation depends on type of data we have

**Observable variables**

Suppose all variables in $X_t, Z_t$ are observable: model makes a clear prediction that given any set of structural parameter $A, B, H, D$, the data will be given by

$$Z_t = CZ_{t-1} + PX_t \tag{4}$$

Likely that there is no set of $A, B, D, H$ matrices that will allow the model to fit the data perfectly

- Cross-equation restrictions in DSGE models
- Particular patterns must be observed by $C, P$
- Cannot use MLE as a result

Can add error terms $u_t$ to estimate $A, B, D, H$

$$Z_t = CZ_{t-1} + PX_t + u_t \tag{5}$$

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

Can use MLE when we have model with observable variables

$$Z_t = CZ_{t-1} + PX_t + u_t \tag{6}$$

$$X_t = DX_{t-1} + \epsilon_t \tag{7}$$

$$u_t \sim N(0, \Sigma_u) \tag{8}$$

$$\epsilon_t \sim N(0, \Sigma_\epsilon) \tag{9}$$

Suppose endogenous and exogenous variables are observed by

$$Z_1, Z_2, ......, Z_T \tag{10}$$

$$X_1, X_2, ......, X_T \tag{11}$$

Can combine log-likelihood fuctions for $Z, X$

- Likelihood full model multiplies likelihood of $X$ data and likelihood of $Z$ data

Maximum likelihood estimates of $A, B, H, D, \Sigma_\epsilon, \Sigma_u$ are those that maximise the following log-likelihood

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

$$
\begin{aligned}
-T\ log 2\pi - T(log|\Sigma_\epsilon^{-1}| + log|\Sigma_u^{-1}|) \\
-\frac{1}{2}\Sigma_{k=1}^{T}(X_i - DX_{i-1})'\Sigma_\epsilon^{-1}(X_i - DX_{i-1}) \\
-\frac{1}{2}\Sigma_{k=1}^{T}(Z_i - CZ_{i-1} - PX_i)'\Sigma_u^{-1}(Z_i - CZ_{i-1} - PX_i)
\end{aligned} \tag{12}
$$

DSGE models mix of observable and unobservable variables

Consider standard RBC model

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

Model features 7 equations: 6 endogenous variables, 1 exogenous variable

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

Model also mixes 4 observable variables and 3 unobservable variables

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

One issue concerning RBC model

- All observed series depend on unobservable technology series
- Same unobserved variable shows up in all reduced-form solution equations

Model has stochastic shocks; also features **stochastic singularity**

- Shocks in equations are multiples of each other
- Model predicts certain rations of observed variables will be constant
    - E.g. current and lagged consumption

Model prediction won't hold in reality; model won't fit data

Implication: every observable variable requires at least one unobservable shock

- Necessary to have well-defined econometric estimates

Unobservable shock can take on two forms:

1. Measurement error
2. Shock with clear structural interpretation

DSGE model, with mix of observable/unobservable variables, is a **state-space** model

1. State (transition) equation $S_t$

$$S_t = FS_{t-1} + u_t$$

2. Measurement equation $Z_t$

$$Z_t = HS_t + v_t$$

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

RBC model (without labour input)

$$k_t = a_{kk} k_{t-1} + a_{kz} z_t \tag{13}$$
$$c_t = a_{ck} k_{t-1} + a_{cz} z_t \tag{14}$$
$$z_t = \rho z_{t-1} + \epsilon_t \tag{15}$$

Assume consumption and capital are only observed with error:

observable variables given by

$$k_t^* = a_{kk} k_{t-1} + a_{kz} z_t + u_t^k \tag{16}$$
$$c_t^* = a_{ck} k_{t-1} + a_{cz} z_t + u_t^c \tag{17}$$

Can rewrite into state-space form: State equation

$$\begin{pmatrix} k_{t-1} \\ z_t \end{pmatrix} = \begin{pmatrix} a_{kk} & a_{kz} \\ 0 & \rho \end{pmatrix} \begin{pmatrix} k_{t-2} \\ z_{t-1} \end{pmatrix} + \begin{pmatrix} 0 \\ \epsilon_t \end{pmatrix} \tag{18}$$

Measurement equation

$$\begin{pmatrix} k^*_{t-1} \\ c^*_t \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ a_{ck} & a_{cz} \end{pmatrix} \begin{pmatrix} k_{t-1} \\ z_t \end{pmatrix} + \begin{pmatrix} u^k_{t-1} \\ u^c_t \end{pmatrix} \tag{19}$$

NB - 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.

System given by

$$S_t = \begin{pmatrix} k_{t-1} \\ z_t \end{pmatrix} \tag{20}$$

$$Z_t = \begin{pmatrix} k^*_{t-1} \\ c^*_t \end{pmatrix} \tag{21}$$

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

Specified model can be estimated on computer which will

1. Sort the model into space-state model
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)

MLE estimation not without dangers:

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

*maximizing a complicated, highly dimensional function like the likelihood of a DSGE model is actually much harder than it is to integrate it, which is what we do in a Bayesian exercise. First, the likelihood of DSGE models is, as I have just mentioned, a highly dimensional object, with a dozen or so parameters in the simplest cases to close to a hundred in some of the richest models in the literature. Any search in a high dimensional function is fraught with peril. More pointedly, likelihoods of DSGE models are full of local maxima and minima and of nearly flat surfaces. This is due both to the sparsity of the data (quarterly data do not give us the luxury of many observations that micro panels provide) and to the flexibility of DSGE models in generating similar behavior with relatively different combination of parameter values .... Moreover, the standard errors of the estimates are notoriously difficult to compute and their asymptotic distribution a poor approximation to the small sample one*

From *"The econometrics of DSGE models"*, Fernandez-Villaverde

Most research uses Bayesian methods

1. Specify prior
2. Combine with likelihood function
3. Calculate posterior estimate
4. Use posterior to produce means, uncertainty intervals, etc.

Main advantage: Get full likelihood function rather than single point estimate

First: data

$$y^T \equiv \{y_t\}_{t=1}^T \in \mathbb{R}^{N \times T} \tag{22}$$

Second: model

$$i \in M \tag{23}$$

Model composed by

1. Parameter set

$$\Theta_i \in \mathbb{R}^{k_i} \tag{24}$$

2. Likelihood function

$$p(y^T|\theta, i) : \mathbb{R}^{N \times T} \times \Theta_i \to \mathbb{R}^+ \tag{25}$$

3. Prior distribution

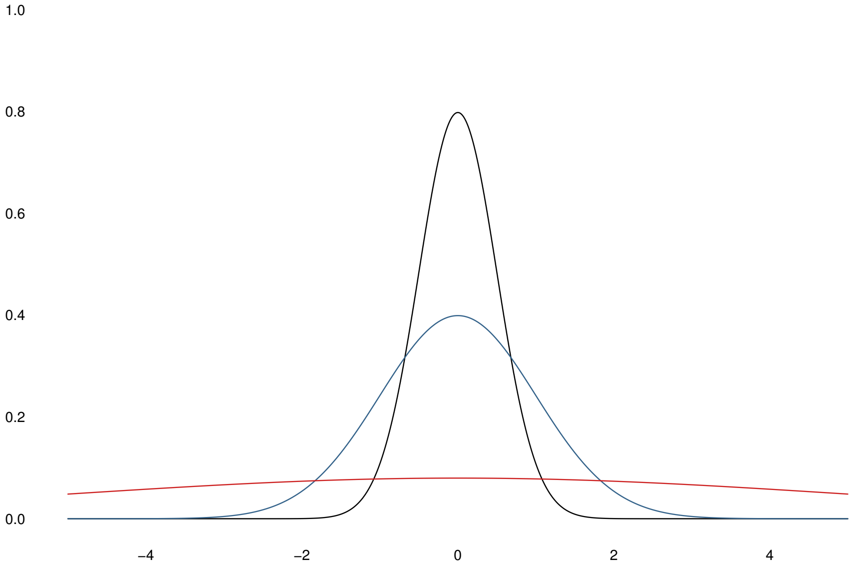
$$\pi(\theta|i) : \Theta_i \to \mathbb{R}^+ \tag{26}$$

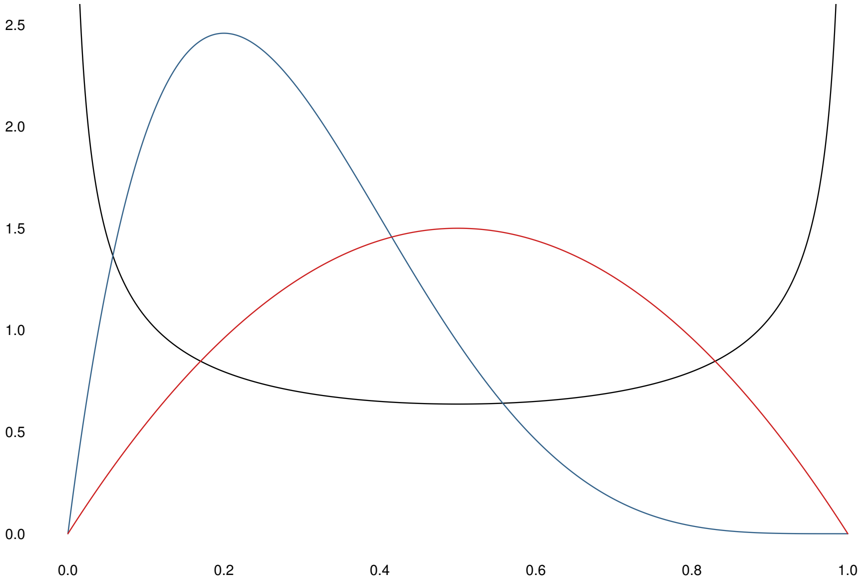
Posterior distribution of parameters given by

$$\pi(\theta|y^T, i) = \frac{p(y^T|\theta, i)\pi(\theta|i)}{\int p(y^T|\theta, i)\pi(\theta|i)d\theta}. \tag{27}$$

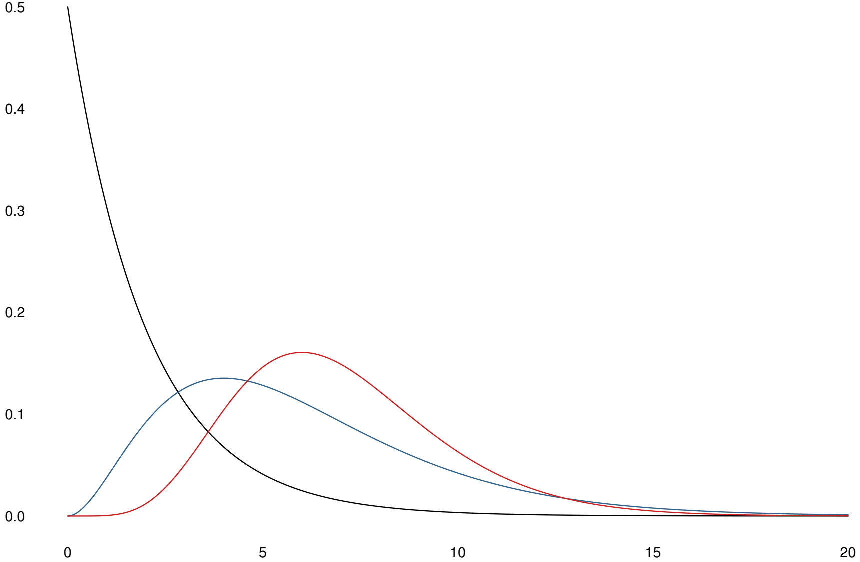
In short: Prior beliefs $\pi(\theta|i)$ are combined with sample information $f(y^T|\theta, i)$ to arrive at new beliefs $\pi(\theta|y^T, i)$

**Q:** How is the prior determined?


1.0
0.8
0.6
0.4
0.2
0.0
−4
−2
0
2
4



2.5
2.0
1.5
1.0
0.5
0.0
0.0
0.2
0.4
0.6
0.8
1.0


0.5
0.4
0.3
0.2
0.1
0.0
0
5
10
15
20

Advantages Bayesian method

1. User-relevant answers
2. Can use pre-sample information
3. Direct computation of objects of interests
4. Ease to deal with misspecified models

Disadvantages Bayesian method

1. Other methods provide more transparent link with FOCs and equilibrium equations
2. Still can be computationally intensive

Need right set of tools to estimate model

1. Solution methods
2. Method to evaluate likelihood of model
3. Method to explore likelihood of model

No analytical solutions

- Need numerical approximations

1. Substitute difficult original problem with simpler one
2. Use solution to approximate solution of original problem

In DSGE: Taylor expansion of function describing variable dynamics around deterministic steady-state

**State-space model**

1. Transition equation

$$S_t = f(S_{t-1}, W_t; \theta)$$

2. Measurement equation

$$Y_t = g(S_t, V_t; \theta)$$

Measurement equation subject to one restriction

1. Can only select number of series less or equal than number of shocks $(W_t, V_t)$

Otherwise it will be stochastically singular: likelihood $-\infty$ with probability 1

Compute

$$S_t = f(S_{t-1}, W_t; \theta) \to p(S_t | S_{t-1}; \theta) \tag{28}$$

$$Y_t = g(S_t, V_t; \theta) \to p(Y_t | S_t; \theta) \tag{29}$$

$$Y_t = g(f(S_{t-1}, W_t; \theta), V_t; \theta) \to p(Y_t | S_{t-1}; \theta) \tag{30}$$

$$p(Y^T|\theta) = p(y_1|\theta)\prod_{t=2}^{T} p(y_t|y_{t-1};\theta) \tag{31}$$

$$= \int (p(y_1|s_1;\theta)dS_1 \prod_{t=2}^{T} \int p(y_t|S_t;\theta)p(S_t|Y_{t-1};\theta)dS_t \tag{32}$$

Can therefore evaluate likelihood of model when we have knowledge of

$$\{p(S_t|y_{t-1};\theta)\}_{t-1}^{T} \tag{33}$$

$$p(S_1;\theta) \tag{34}$$

**Chapman-Kolmogorov equation**

$$p(S_{t+1}|y_t;\theta) = \int p(S_{t+1}|S_t;\theta)p(S_t|y_t;\theta)dS_t \tag{35}$$

Provides forecasting rule for the evolution of states

> *The distribution of states tomorrow, given the states until today, is equal to the distribution of states today times the transition probabilities integrated over all possible states*

**Bayes theorem** once again

$$p(S_t|y_t;\theta) = \frac{p(y_t|S_t;\theta)p(S_t|y_{t-1};\theta)}{p(y_t|y_{t-1};\theta)} \tag{36}$$

$$p(y_t|y_{t-1};\theta) = \int p(y_t|S_t;\theta)p(S_t|y_{t-1};\theta)dS_t \tag{37}$$

Using Chapman-Kolmogorov equation and Bayes Theorem can generate complete sequence for

$$\{p(S_t|y_{t-1};\theta)\}_{t=1}^{T} \tag{38}$$

Mathematically straightforward; practically difficult

Can fix computational problem

1. Kalman filter
2. Particle filter
    - Use when state-space representation is not linear or shocks are not normal

**Exploring the likelihood function**

1. Maximisation
2. Description

Use posterior distribution

$$\pi(\theta|y_t) = \frac{p(y_t|\theta)\pi(\theta)}{\int p(y_t|\theta)\pi(\theta)d\theta} \tag{39}$$

Use Markov Chain Monte Carlo (MCMC) methods

- Metropolis-Hastings algorithm
- Gibbs sample (special case of Metropolis-Hastings)

## Metropolis-Hastings algorithm

---

**Step 0, Initialization:** Set $i \rightsquigarrow 0$ and an initial $\theta_i$. Solve the model for $\theta_i$ and build the state space representation. Evaluate $\pi(\theta_i)$ and $p(y^T|\theta_i)$. Set $i \rightsquigarrow i+1$.

**Step 1, Proposal draw:** Get a draw $\theta_i^*$ from a proposal density $q(\theta_{i-1}, \theta_i^*)$.

**Step 2, Solving the Model:** Solve the model for $\theta_i^*$ and build the new state space representation.

**Step 3, Evaluating the proposal:** Evaluate $\pi(\theta_i^*)$ and $p(y^T|\theta_i^*)$ with (9).

**Step 4, Accept/Reject:** Draw $\chi_i \sim U(0,1)$. If $\chi_i \leq \frac{p(y^T|\theta)\pi(\theta)q(\theta_{i-1},\theta_i^*)}{p(y^T|\theta)\pi(\theta)q(\theta_i^*,\theta_{i-1})}$ set $\theta_i = \theta_i^*$, otherwise $\theta_i = \theta_{i-1}$.

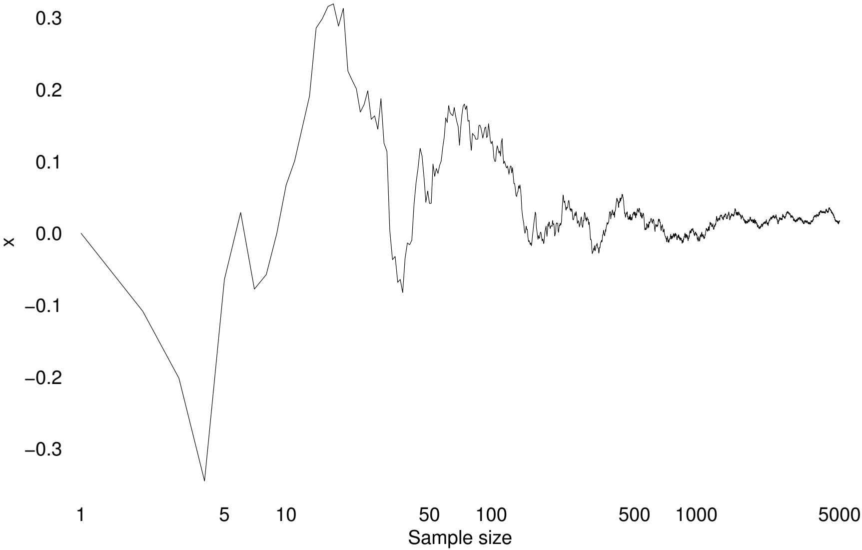
**Step 5, Iteration:** If $i < M$, set $i \rightsquigarrow i+1$ and go to step 1. Otherwise stop.

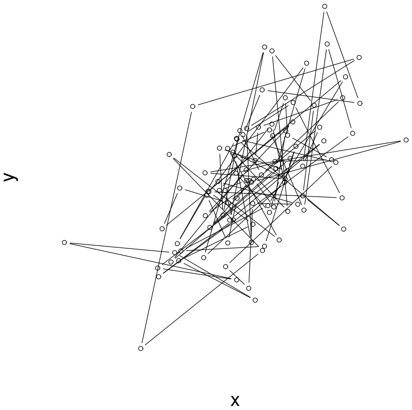
---

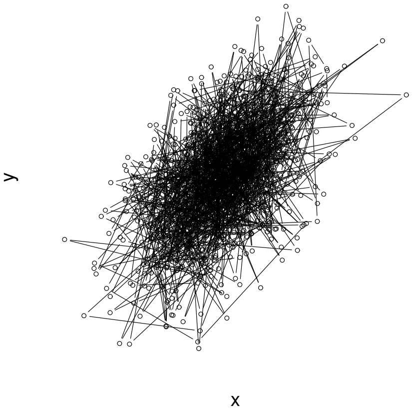
Eq. (9) as referred to in previous slide

$$p(y^T|\theta) \simeq \frac{1}{N}\sum_{i=1}^{N} p(y_1|s_{0|0}^i;\psi)\prod_{t=2}^{T}\frac{1}{N}\sum_{i=1}^{N} p(y_t|s_{t|t-1}^i;\theta)$$

From *"The econometrics of DSGE models"*, Fernandez-Villaverde


0.3
0.2
0.1
0.0
–0.1
–0.2
–0.3
x
1
5
10
50
100
500
1000
5000
Sample size


y
x


y
x