The Econometrics of DSGE Models

Giuseppe Ragusa
Luiss University

EIEF
Lecture 1: Introduction

May 10, 2013

Outline of the course

1. Motivation: DSGE models and their applications
2. Approximating and solving DSGE models
    1. State space representation
    2. Constructing log-linear approximation
    3. Higher order approximation techniques
3. Frequentist estimation of DSGE models
4. Bayesian estimation of DSGE models
5. The twilight zone of DSGE estimation

Course material

Textbook:

Canova, Fabio. Methods for applied macroeconomic research. Vol. 13. Princeton University Press, 2011.

Papers:

An, Sungbae, and Frank Schorfheide. "Bayesian analysis of DSGE models." Econometric reviews 26.2-4 (2007): 113-172.

More introduced as we move along

Slides:

Will make them available @ www.gragusa.org/eief_dsge

Evaluation

1. Final Test (50%)
   $\mapsto$Theoretical questions

2. Problem sets (50%)
   $\mapsto$Based on real world example (Matlab)
   $\mapsto$ Assigned in class and due in one week

DSGE quick history

Kydland, Finn E., and Edward C. Prescott. "Time to build and aggregate fluctuations." Econometrica: Journal of the Econometric Society (1982): 1345-1370.

- For the first time, macroeconomists had a small and coherent dynamic model of the economy, built from first principles with optimizing agents, rational expectations, and market clearing, that could generate data that resembled observed variables to a remarkable degree.
- There were many dimensions along which the model failed, from the volatility of hours to the persistence of output.
- But the amazing feature was how well the model did despite having so little of what was traditionally thought of as the necessary ingredients of business cycle theories: money, nominal rigidities, or non-market clearing.

DSGE quick history, ctd.

- The initial reaction to Kydland and Prescott's assertions varied from incredulity to straightforward dismissal
- Reasons? Due to the main characteristics:
  - The efficiency of business cycles:
    - The bulk of economic fluctuations observed in industrialized countries could be interpreted as an equilibrium outcome resulting from the economy's response to exogenous variations in real forces in an environment with perfect competition and frictionless markets
  - Importance of technological shocks
    - Basic model able to generate "realistic" fluctuations in output and other macroeconomic variables
  - The limited role of monetary factors
    - Theory sought to explain economic fluctuations with no reference to monetary factors, even abstracting from the existence of a monetary sector
  - Technical complexity

DSGE quick history, ctd.

- Nowadays, DSGE model—leveraging on the initial intuition of Kydland and Prescott's (1982) paper—are flexible enough to accommodate rigidities
  - Monopolistic competition: Without market power, any firm that does not immediately adjust its prices will lose all its sales; (Blanchard and Kiyotaki, 1987)
  - Nominal rigidities: firms are subject to constraints on the frequency with which they can adjust the prices of goods and services they sell (Calvo, 1983)
  - Short run non-neutrality of monetary policy: A monetary policy rule, such as a money growth process or a Taylor rule, is used as a monetary authority.
- Reference:
  - Woodford, Michael. Interest and prices: Foundations of a theory of monetary policy. princeton university press, 2011.

DSGE Model

- We will now introduce a simple DSGE model that does not have all the features that I have just enumerated
- This is for two reasons:
  - I need to keep the presentation of the material accessible—realistic models (Smets and Wouters, 2003) are too large to be used in a didactic setting
  - Fortunately, we do not loose much. In fact, considering a prototypical DSGE model will help us in highlighting the econometric issues that arise in a realistic settings without having to deal with theirs complications

DSGE Models

General representation

$$E_t\left[f(\zeta_{t+1},\zeta_t,\zeta_{t-1},\varepsilon_t,\theta)\right]=0$$

- Collection of agents' first order conditions and constraints
- Notation
  - ζ_t : vector of endogenous variables
  - ε_t : vector of structural shocks, (e.g., $\log\varepsilon_t \sim i.i.d.\, N(0,\Sigma)$)
  - θ : vector of structural parameters

DSGE Models

General representation

$$E_t \left[f(\zeta_{t+1}, \zeta_t, \zeta_{t-1}, \varepsilon_t, \theta) \right] = 0$$

- Collection of agents' first order conditions and constraints
- Notation
  - ζ_t : vector of endogenous variables
  - ε_t : vector of structural shocks, (e.g., $\log \varepsilon_t \sim i.i.d. N(0, \Sigma)$)
  - θ : vector of structural parameters

DSGE Models

General representation

$$E_t\left[f(\zeta_{t+1}, \zeta_t, \zeta_{t-1}, \varepsilon_t, \theta)\right] = 0$$

- Collection of agents' first order conditions and constraints
- Notation
  - ζ_t : vector of endogenous variables
  - ε_t : vector of structural shocks, (e.g., $\log \varepsilon_t \sim i.i.d.\,N(0, \Sigma)$)
  - θ : vector of structural parameters

DSGE Models

General representation

$$E_t \left[f(\zeta_{t+1}, \zeta_t, \zeta_{t-1}, \varepsilon_t, \theta) \right] = 0$$

- Collection of agents' first order conditions and constraints
- Notation
  - ζ_t : vector of endogenous variables
  - ε_t : vector of structural shocks, (e.g., $\log \varepsilon_t \sim i.i.d. N(0, \Sigma)$)
  - θ : vector of structural parameters

Neoclassical Growth Model

Setup

$$U = E_0 \sum_{t=1}^{\infty} \beta^{t-1} \frac{\left(C_t^{\lambda}(1-H_t)^{1-\lambda}\right)^{1-\tau}}{1-\tau}$$

$$Y_t = C_t + I_t$$

$$Y_t = e^{z_t} K_t^{\alpha} H_t^{1-\alpha}$$

$$K_{t+1} = I_t + (1-\delta)K_t$$

$$z_t = \rho z_{t-1} + \varepsilon_t, \ |\rho| < 1$$

$$\varepsilon_t \sim N(0, \sigma_\varepsilon^2)$$

- C_t : consumption,
- H_t : hours
- Y_t : product
- K_t : capital
- I_t : investment
- z_t: technology shocks
- u_t: exogenous shock
- $\underbrace{\theta = (\beta, \lambda, \tau, \alpha, \delta, \rho, \sigma_u^2)}_{\text{structural parameters}}$

Neoclassical Growth Model

- Both welfare theorems hold in this economy.
- Thus, we can solve directly for the social planner's problem:

$$\max_{\{C_t,H_t\}} E_0 \sum_{t=1}^{\infty} \beta^{t-1} \frac{\left(C_t^{\lambda}(1-H_t)^{1-\lambda}\right)^{1-\tau}}{1-\tau}$$

subject to

$$C_t + I_t = e^{z_t} K_t^{\alpha} H_t^{1-\alpha}$$

$$K_{t+1} = I_t + (1-\delta) K_t$$

$$z_t = \rho z_{t-1} + \varepsilon_t,\ |\rho| < 1,\ \varepsilon_t \sim N(0, \sigma_{\varepsilon}^2)$$

$$z_0, K_0$$

  - maximize the utility of the household subject to the production function, the evolution of technology, the law of motion for capital, the resource constraint, and some initial k_0 and z_0.

Neoclassical Growth Model

First Order Conditions

The model is fully characterized by the first order conditions:

$$1 = E_t \left[\beta \left(\frac{C_{t+1}}{C_t} \right)^{\lambda(1-\tau)} \left(\frac{1 - H_{t+1}}{1 - H_t} \right)^{(1-\lambda)(1-\tau)} R_{t+1} \right]$$

$$R_{t+1} \equiv (1-\delta) + \alpha e^{z_{t+1}} K_{t+1}^{\alpha-1} H_{t+1}^{1-\alpha}$$

$$(1-\lambda) \frac{1}{(1-H_t)} = \frac{\lambda(1-\alpha) e^{z_t} K_t^{\alpha} H_t^{-\alpha}}{C_t}$$

$$C_t + K_{t+1} = e^{z_t} K_t^{\alpha} H_t^{1-\alpha} + (1-\delta) K_t$$

$$z_t = \rho z_{t-1} + \varepsilon_t.$$

It "clearly" fits

$$E_t \left[f(\zeta_{t+1}, \zeta_t, \zeta_{t-1}, \varepsilon_t, \theta) \right] = 0$$

DSGE Models

Solution

Policy function

The solution of the dynamic system is a decision rule

$$\underbrace{\zeta_t = r(\zeta_{t-1}, \varepsilon_t; \theta)}_{\text{policy function}}$$

- Policy function are not available in closed form (except for very stylized models)
- We often forced to resort to numerical approximations to characterize the equilibrium dynamics of the model
- Three "historical" approaches
  - using a linear quadratic approximation of the model (Kydland and Prescott, 1982)
  - value function iterations (Christiano, 1990)
  - (log-)linearization of equilibrium conditions (King, Plosser, Rebelo, 2002)

In this course we will focus on the last approach, since it is the one used in

DSGE Model

Solution: Linearization

The three main ingredients are:

1. Non-stochastic steady state (NSSS)
2. Linearization
3. Gensys' canonical form

DSGE Model

Solution: Linearization

The three main ingredients are:

1. Non-stochastic steady state (NSSS)
2. Linearization
3. Gensys' canonical form

The three main ingredients are:

1. Non-stochastic steady state (NSSS)
2. Linearization
3. Gensys' canonical form

DSGE Models

Non-stochastic steady state

The NSSS of the model is found by

1. setting the shock to $z_t = \bar{z} = 0$
2. dropping the time indexes
3. solving for the endogenous variables.

Neoclassical Growth Model

Linearization

For the Neoclassical Growth Model, the NSSS is given:

$$K_{ss} = S_1/(S_2 S_1 + S_3) \quad (1)$$

$$H_{ss} = S_2 K_{ss} \quad (2)$$

$$C_{ss} = S_3 K_{ss} \quad (3)$$

$$Y_{ss} = K_{ss}^{\alpha} H_{ss}^{1-\alpha}, \quad (4)$$

where

$$S_1 = \frac{\lambda(1-\alpha)S_2^{-\alpha}}{(1-\lambda)}, \quad S_2 = \frac{(1/\beta - 1 + \delta))^{1/(1-\alpha)}}{\alpha}, \quad S_3 = (S_2^{1-\alpha} - \delta).$$

Note: The NSSS is a function of the parameters.

DSGE Model

Linearization

The linearization gives a set of linear equations that describe the model around the NSSS.

$$\Gamma_0(\theta)E[x_{t+1}] = \Gamma_1 x_t(\theta) + C + \Psi(\theta)z_t$$

- The variables (x_t) are expressed in deviation from the NSSS
- $\Gamma_0(\theta)$, $\Gamma_1(\theta)$, and $\Psi(\theta)$ are function of the structural parameter θ
- Several solution algorithms are available for this linear rational expectations system: e.g.,
  - Blanchard, Olivier Jean, and Charles M. Kahn. "The solution of linear difference models under rational expectations." Econometrica: Journal of the Econometric Society (1980): 1305-1311.
  - Sims, Christopher A. "Solving linear rational expectations models." Computational Economics 20.1 (2002): 1-20.
- All this different linear methods are equivalent, as the linear approximation of a differentiable function is unique and invariant (differentiable parameters transformations)

Neoclassical Growth Model

Linearization

For $k_t = K_t - K_{ss}$ and $h_t = H_t - H_{ss}$, the linearized version of the model is

$$Ak_{t+1} + Bk_t + Ch_t + Dz_t = 0 \tag{5}$$

$$E_t \left(Gk_{t+1} + Hk_t + Jh_{t+1} + Kh_t + Lz_{t+1} + Mz_t \right) = 0 \tag{6}$$

$$Ez_{t+1} = Nz_t, \tag{7}$$

where the coefficient $A, B, C, \ldots, N$ are functions of the structural parameter θ.

Neoclassical Growth Model

Linearization, ctd.

Expression for the coefficient of the policy functions

$$
\begin{aligned}
A &= 1\\
B &= (\alpha/K_{ss})C_{ss} - Y_{ss}(\alpha/K_{ss}) - (1-\delta)\\
C &= -((\alpha/H_{ss}) + 1/(1-H_{ss}))C_{ss} - Y_{ss}((1-\alpha)/H_{ss})\\
D &= C_{ss} - Y_{ss}\\
G &= \alpha(\lambda(1-\tau)-1)/K_{ss}\\
H &= (\lambda(1-\tau)-1)/K_{ss}\\
J &= -\alpha(\lambda(1-\tau)-1)((\alpha/H_{ss}) + (1/(1-H_{ss})))\\
&\quad + \beta(\alpha(1-\alpha)/H_{ss})K_{ss}^{\alpha-1}H_{ss}^{1-\alpha} - ((1-\tau)(1-\lambda)/(1-H_{ss}))\\
K &= \alpha(\lambda(1-\tau)-1)((\alpha/H_{ss}) + (1/(1-H_{ss}))) - (1-\lambda)(1-\tau)/(1-H_{ss})\\
L &= \lambda(1-\tau) - H_{ss} + \alpha\beta K_{ss}^{\alpha-1}H_{ss}^{1-\alpha}\\
M &= \lambda(1-\tau) - H_{ss}\\
N &= \rho
\end{aligned}
$$

Neoclassical Growth Model

Solution

The solution of the system of rational expectation equation is given by:

$$
\begin{aligned}
k_{t+1} &= P(\theta)k_t + Q(\theta)z_t \\
h_{t+1} &= R(\theta)\,P(\theta)k_t + (R(\theta)\,Q(\theta) - S(\theta)\rho)z_t + S(\theta)\varepsilon_{t+1} \\
z_{t+1} &= \rho z_t + \varepsilon_{t+1}.
\end{aligned}
$$

We can thus write

$$
x_{t+1} = G_0(\theta) + G(\theta)x_t + M(\theta)\varepsilon_{t+1}
$$

where

$$
x_{t+1} = \begin{bmatrix} k_{t+1} \\ h_{t+1} \\ z_{t+1} \end{bmatrix}, \quad G(\theta) = \begin{bmatrix} P(\theta) & 0 & Q(\theta) \\ R(\theta)\,P(\theta) & 0 & R(\theta)\,Q(\theta) - S(\theta)\rho \\ 0 & 0 & \rho \end{bmatrix}, \quad M(\theta
$$

Neoclassical Growth Model

Solution, ctd.

The coefficients of the policy functions are:

$$P = \frac{1}{2}\left(\frac{G\,C}{J\,A} - \frac{B}{A} - \frac{K}{J} - \sqrt{\left(\frac{\frac{B}{A} + \frac{K}{J} - \frac{G\,C}{J\,A}}{(J\,A)}\right)^2 - 4\frac{K\,B - H\,C}{J\,A}}\right)$$

$$R = -\frac{1}{C}(A\,P\,B)$$

$$Q = \frac{C\,L\,N - D(J + N + K)}{A\,J\,N + A\,K - C\,G - C\,J\,R}$$

$$S = \frac{D\,G + D\,J\,R - A\,L\,N - A\,M}{A\,J\,N + A\,K - C\,G - C\,J\,R}$$

DSGE Models

1. Non-stochastic steady state (NSSS)
2. Linearization
3. Gensys' canonical form

$$\Downarrow$$

$$x_{t+1} = G_0(\theta) + G(\theta)x_t + M(\theta)\varepsilon_{t+1}$$

- State/transition equation
- Describe the evolution of the model's endogenous variables
- Data???????

DSGE Models

1. Non-stochastic steady state (NSSS)
2. Linearization
3. Gensys' canonical form

$$\Downarrow$$

$$x_{t+1} = G_0(\theta) + G(\theta)x_t + M(\theta)\varepsilon_{t+1}$$

- State/transition equation
- Describe the evolution of the model's endogenous variables
- Data???????

DSGE Models

1. Non-stochastic steady state (NSSS)
2. Linearization
3. Gensys' canonical form

$$\Downarrow$$

$$x_{t+1} = G_0(\theta) + G(\theta)x_t + M(\theta)\varepsilon_{t+1}$$

- State/transition equation
- Describe the evolution of the model's endogenous variables
- Data???????

DSGE Models

State-space representation

The data enter through the specification of a set of measurement equations:

$$\underbrace{\underbrace{y_t}_{\text{observables}} = H(\theta)x_t \underbrace{(+m(\theta)\eta_t)}_{\text{meas. error}}}_{\text{observation equation}}$$

State-space representation

The reduced-form model is given by the following state-space form:

$$x_{t+1} = G_0(\theta) + G(\theta)x_t + M(\theta)\varepsilon_{t+1}$$

$$y_t = H(\theta)x_t \,(+m(\theta)\eta_{t+1})$$

DSGE Models

State-space representation

The data enter through the specification of a set of measurement equations:

$$\underbrace{\underbrace{y_t}_{\text{observables}} = H(\theta)x_t \underbrace{(+m(\theta)\eta_t)}_{\text{meas. error}}}_{\text{observation equation}}$$

State-space representation

The reduced-form model is given by the following state-space form:

$$x_{t+1} = G_0(\theta) + G(\theta)x_t + M(\theta)\varepsilon_{t+1}$$

$$y_t = H(\theta)x_t\,(+m(\theta)\eta_{t+1})$$

DSGE Models

Solutions: Remarks

- The "true" model is highly non-linear:

$$E_t\left[f(\zeta_{t+1},\zeta_t,\zeta_{t-1},\varepsilon_t,\theta)\right]=0$$

- We end up with a linear model:

$$\begin{aligned} x_{t+1} &= G_0(\theta)+G(\theta)x_t+M(\theta)\varepsilon_{t+1} \\ y_{t+1} &= H(\theta)x_t\,(+m(\theta)\eta_t) \end{aligned}$$

- linearization may have important implications
- ....but it simplifies estimation of the model.

DSGE Models

Solutions: Remarks

- The "true" model is highly non-linear:

$$E_t\left[f(\zeta_{t+1}, \zeta_t, \zeta_{t-1}, \varepsilon_t, \theta)\right] = 0$$

- We end up with a linear model:

$$\begin{aligned} x_{t+1} &= G_0(\theta) + G(\theta)x_t + M(\theta)\varepsilon_{t+1} \\ y_{t+1} &= H(\theta)x_t\,(+m(\theta)\eta_t) \end{aligned}$$

- linearization may have important implications
- ....but it simplifies estimation of the model.

DSGE Models

Solutions: Remarks

- The "true" model is highly non-linear:

$$E_t\left[f(\zeta_{t+1}, \zeta_t, \zeta_{t-1}, \varepsilon_t, \theta)\right] = 0$$

- We end up with a linear model:

$$\begin{aligned} x_{t+1} &= G_0(\theta) + G(\theta)x_t + M(\theta)\varepsilon_{t+1} \\ y_{t+1} &= H(\theta)x_t \,(+m(\theta)\eta_t) \end{aligned}$$

- linearization may have important implications
- ....but it simplifies estimation of the model.

DSGE Models

Solutions: Remarks

- The "true" model is highly non-linear:

$$E_t\left[f(\zeta_{t+1}, \zeta_t, \zeta_{t-1}, \varepsilon_t, \theta)\right] = 0$$

- We end up with a linear model:

$$\begin{aligned} x_{t+1} &= G_0(\theta) + G(\theta)x_t + M(\theta)\varepsilon_{t+1} \\ y_{t+1} &= H(\theta)x_t\,(+m(\theta)\eta_t) \end{aligned}$$

- linearization may have important implications
- ....but it simplifies estimation of the model.

Estimation.....

Given the state space model:

$$x_{t+1} = G_0(\theta) + G(\theta)x_t + M(\theta)\varepsilon_{t+1}$$
$$y_t = H(\theta)x_t\,(+m(\theta)\eta_t)$$

and knowing θ, the structural parameter vector, we can—using standard filtering techniques fit the model to the data and assess its goodness of fit

- This is the approach that we will explore next,
- ...but first let's focus on the "glorious" past

GMM estimation

- First attempt to estimate parameters of macroeconomic models where limited to subset of θ
- Consider the Euler conditions of the neoclassical growth model

$$1 = E_t\left[\beta\left(\frac{C_{t+1}}{C_t}\right)^{\lambda(1-\tau)}\left(\frac{1-H_{t+1}}{1-H_t}\right)^{(1-\lambda)(1-\tau)} R_{t+1}\right]$$

- For simplicity let $\lambda = 1$, so that it reduces to

$$1 = E_t\left[\beta\left(\frac{C_{t+1}}{C_t}\right)^{-\alpha} R_{t+1}\right], \quad \alpha \equiv \tau - 1.$$

  The objective is to estimate (β, α).
- Hansen and Singleton (1982) do that building on Hansen (1982) celebrated GMM paper.

GMM Estimation

- Notice that the expectation is conditional to the information available at time t
- By the law of iterated expectation

$$E_t\left[\beta\left(\frac{C_{t+1}}{C_t}\right)^{-\alpha}R_{t+1}-1\right]=0$$

implies

$$E\left[\left(\beta\left(\frac{C_{t+1}}{C_t}\right)^{-\alpha}R_{t+1}-1\right)z_t\right]=0,$$

for any vector of variables z_t that is known at time t.

GMM Estimation

- Having observation on consumption and interest rates we could estimate (β, α) by solving the empirical moment equation

$$\frac{1}{T}\sum_{t=1}^{T}\left(\beta\left(\frac{C_{t+1}}{C_t}\right)^{-\alpha}R_{t+1}-1\right)z_t=0.$$

- Two cases:
  - $dim(z_t)=2$: the system is exactly identified and thus is reasonable that a solution exists
  - $dim(z_t)>2$: the system is overidentified, thus no solution exists
- Assume that $dim(z_t)=m\geq 2$.

GMM Estimation

Let

$$g(w_t, \theta) \equiv \left(\beta \left(\frac{C_t}{C_{t-1}} \right)^{-\alpha} R_t - 1 \right) z_t,$$

where $w_{t+1} = \{C_{t+1}, C_t, R_{t+1}\}$ and $\theta = \{\beta, \alpha\}$.
The GMM estimator is given by

$$\hat{\theta}^{GMM} = \arg\min_{\theta \in \Theta} \left[\underbrace{\frac{1}{T} \sum_{t=1}^{T} g(w_t, \theta)'}_{1 \times m} \right] \underbrace{W}_{m \times m} \left[\underbrace{\frac{1}{T} \sum_{t=1}^{T} g(w_t, \theta)}_{m \times 1} \right],$$

where W is a positive definite matrix.

GMM Estimation

Let θ_0 denote the "true" parameters, that is,

$$E[g(w_t, \theta_0)] = 0.$$

Under regularity conditions,

$$\hat{\theta} \xrightarrow{p} \theta_0$$

and

$$\sqrt{T}(\hat{\theta} - \theta_0) \xrightarrow{d} N(0, V_W)$$

- The GMM estimator is consistent regardless of the weighting matrix
- The GMM estimator is asymptotically normal with a variance that depends on the weighting matrix W, V_W
- Which W should we use? ...the one that minimizes the asymptotic variance.....

GMM Estimation

The (optimal, efficient) GMM estimator is defined as

$$\hat{\theta}^{GMM} = \arg\min_{\theta\in\Theta} \left[\underbrace{\frac{1}{T}\sum_{t=1}^{T} g(w_t,\theta)'}_{1\times m} \right] \underbrace{\Omega^{-1}}_{m\times m} \left[\underbrace{\frac{1}{T}\sum_{t=1}^{T} g(w_t,\theta)}_{m\times 1} \right],$$

where

$$\Omega = E[g(w_t,\theta_0)g(w_t,\theta_0)'].$$

In this case,

$$\sqrt{T}(\hat{\theta}^{GMM} - \theta) \xrightarrow{d} N(0,V),\ V = (G'\Omega^{-1}G)^{-1},\ G = E[\partial g(w_t,\theta_0)/\partial\theta_0]$$

- Hansen (1982) shows that $V < V_W$.

Two-step procedures

- In practice Ω is not know
- Can be estimated using a preliminary consistent estimator of θ_0, say $\tilde{\theta} \xrightarrow{p} \theta_0$

$$\widehat{\Omega} = \frac{1}{T} \sum_{t=1}^{T} g(w_t, \tilde{\theta}) g(w_t, \tilde{\theta})'$$

- Then solve

$$\hat{\theta}^{GMM} = \arg\min_{\theta \in \Theta} \left[\underbrace{\frac{1}{T} \sum_{t=1}^{T} g(w_t, \theta)'}_{1 \times m} \right] \underbrace{\widehat{\Omega}^{-1}}_{m \times m} \left[\underbrace{\frac{1}{T} \sum_{t=1}^{T} g(w_t, \theta)}_{m \times 1} \right]$$

- It is a two step procedures because $\tilde{\theta}$ is a GMM with non-optimal weighting matrix

Test of overidentified restrictions

- If the model is correctly specified, that is, there exists a θ_0 such that

$$E[g(w_t, \theta_0)] = 0,$$

  then

$$J = T \left[\underbrace{\frac{1}{T} \sum_{t=1}^{T} g(w_t, \hat{\theta}^{GMM})'}_{1 \times m} \right] \underbrace{\widehat{\Omega}^{-1}}_{m \times m} \left[\underbrace{\frac{1}{T} \sum_{t=1}^{T} g(w_t, \hat{\theta}^{GMM})}_{m \times 1} \right] \xrightarrow{d} \chi^2_{m-k}$$

- J statistics—as it is commonly referred to—can then be used to test whether the model is correctly specified (but in reality we are testing whether the over-identified restrictions are satisfied)

Application to US economy

- Per capita real consumption (non-durables plus services) C_t
- Real risk free rate r_t^{rf}, $R_t^{rf} = 1 + r_t^{rf}$
- Instruments

$$z_t^{(1)} = \begin{pmatrix} 1 \\ C_t / C_{t-1} \end{pmatrix}, \quad z_t^{(2)} = \begin{pmatrix} 1 \\ C_t / C_{t-1} \\ 1 + \pi_t \end{pmatrix}$$

US data

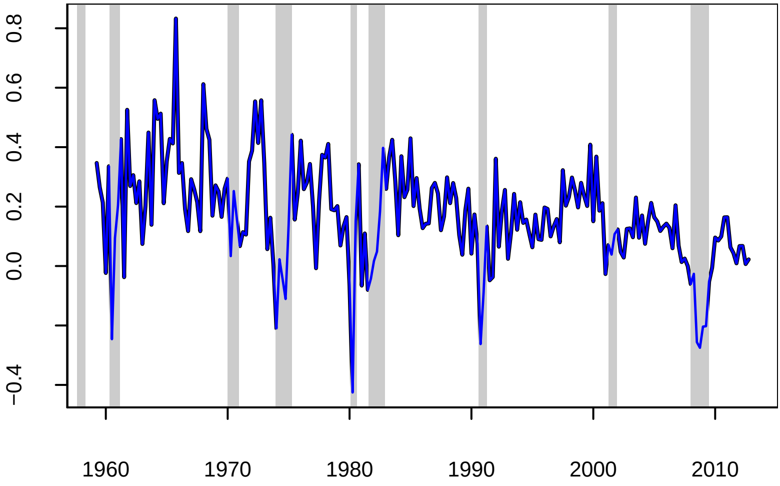

Figure : Per capita real consumption growth, nondurables plus services. Quarterly data: 1959:2-2012:4

US data

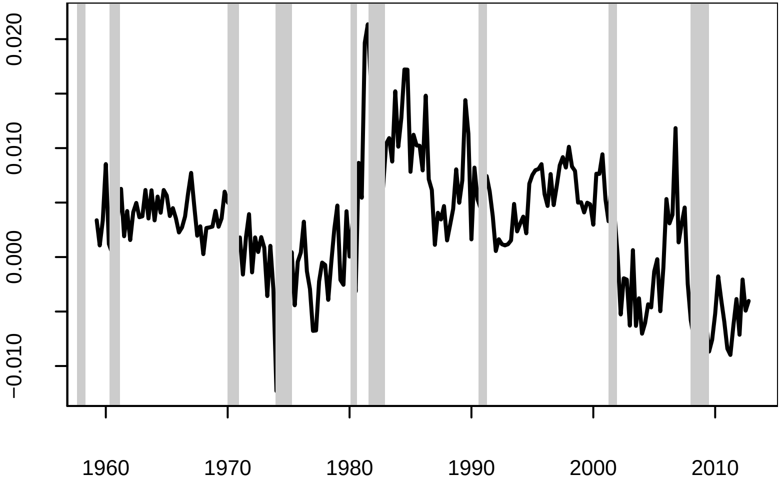

Figure : Real risk free rate. Quarterly data: 1959:2-2012:4

US data

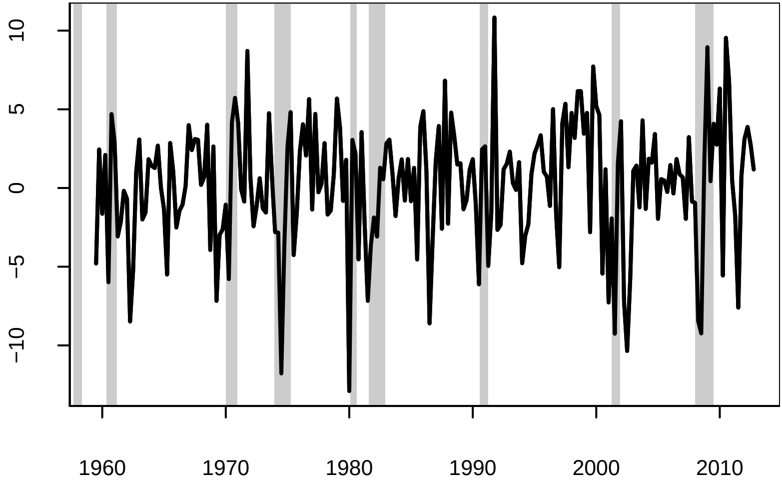

Figure : Excess returns. Quarterly data: 1959:2-2012:4

GMM estimation: Exact identification

```
GMM estimation (# interations: 2)

Coefficients:
       Estimate Std. Error t value Pr(>|z|)
theta1   0.8698     0.5238   1.661   0.0484 *
theta2   0.9985     0.0011 899.984   0.0000 ***

Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1
Objective value: 1.087e-16  on  0  DF, p-value:  < 2.2e-16
Observations: 214, moment restrictions: 2
```

GMM estimation: Over-identification

```
GMM estimation (# interations: 2)

Coefficients:
       Estimate Std. Error  t value Pr(>|z|)
theta1   0.8795     0.4240    2.074    0.019 *
theta2   0.9990     0.0009 1169.603    0.000 ***

Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1
Objective value: 25.97  on  1  DF, p-value:  3.46e-07
Observations: 214, moment restrictions: 3
```

Testing the model

- The J statistic is very large:

$$J = 25.97$$

- Since $m - k = 1$ in this case, the critical value for a 5% significance test is 3.84

$$\Pr(\chi_1^2 > 3.84) \approx 0.05.$$

- Since

$$25.97 > 3.84$$

we reject the null hypothesis the the over-identified conditions hold at the 5% significance level (actually we reject at any reasonable significance level)

GMM critiques

- Small sample performances of GMM are often very poor
  - Special issues of Journal of Business and Economic Statistics: vol. 14, no. 3
  - Hansen, Lars Peter; Heaton, John; Yaron, Yaron (1996). "Finite-sample properties of some alternative GMM estimators". Journal of Business & Economic Statistics 14 (3): 262–280.
- Alternative estimator
  - Continuous updating GMM
  - Generalized Empirical Likelihood (GEL)
    - ★ Imbens, Guido W.; Spady, Richard H.; Johnson, Phillip (1998). "Information theoretic approaches to inference in moment condition models". Econometrica 66 (2): 333–357

Continuous updating

The Continuous Updating GMM estimates minimizes the objective function over θ_0

$$\arg\min_{\theta\in\Theta}\left[\underbrace{\frac{1}{T}\sum_{t=1}^{T}g(w_t,\theta)'}_{1\times m}\right]\underbrace{\left[\frac{1}{T}\sum_{t=1}^{T}g(w_t,\theta)g(w_t,\theta)'\right]^{-1}}_{m\times m}\left[\underbrace{\frac{1}{T}\sum_{t=1}^{T}g(w_t,\theta)}_{m\times 1}\right]$$

- In Monte-Carlo experiments this method demonstrated a better performance than the traditional two-step GMM: the estimator has smaller median bias (although fatter tails), and the J-test for overidentifying restrictions in many cases was more reliable
- Implementation is somewhat more difficult.....

CUE Estimation

```
MD estimation: "Continuous Updating"

Coefficients:
       Estimate Std. Error  t value Pr(>|z|)
theta1   0.8444     0.3794    2.226    0.013 *
theta2   0.9990     0.0008 1186.192    0.000 ***

Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1
Objective value: 29.43  on  1  DF, p-value:  5.801e-08
Observations: 214, moment restrictions: 3, variance: wr
```