University of Dublin,

Trinity College

FACULTY OF ARTS, HUMANITIES AND SOCIAL SCIENCES

Department of Economics

M.Sc. in Economics **Hilary Term 2013**

Econometrics II

(Date) **(Venue)** **(Time)**

Prof Charles Nolan, Prof Agustin Benetrix and Mr Michael Curran

Instructions to candidates:

Please attempt TWO questions from section A and TWO questions from section B.

Please return the exam paper with your scripts.

Materials Permitted for this Examination

Standard non-programmable calculator.

You may not start this examination until you are instructed to do so by the Invigilator.

Section A

Please attempt TWO questions from the three questions in this section.

Question 1 (100 Marks) -- Identification & Frequency Related Filtering.

Part (a): (70 Marks)

Suppose there is an epidemic in Ireland and there are only two possible, mutually exclusive treatments for the disease, drug a or drug b, where say drug a is a status quo treatment and drug b is some innovation. Half the patients in the study population have been treated with $t = a$ and half have been treated with $t = b$. A physician obtains data on these treatment diseases and observes partial data on the number of years Y, that each patient lives after treatment. Table 2 shows the available data on the distribution.

	Treatment		
years of life after treatment	(Z=a)	(Z=b)	total
Y = 0	.10	.12	.22
Y = 1	.25	.30	.55
Y = 2	.15	.08	.23
total	.50	.50	1

Table 2: Treatment under ambiguity.

i. Given the available data, what can the physician deduce about $P[Y(a) = 0]$? Show your workings.
ii. Given the available data, what can the physician deduce about the average treatment effect, $E[Y(b)] - E[Y(a)]$? Show your workings.
iii. What is the maximin treatment rule? What is the minimax-regret treatment rule? Show your workings.
iv. Let x be a dummy variable indicating membership of a Jazz band. Observing $P(y = 2 \mid x = 0, z = a) = 0.25$ and $P(y = 2 \mid x = 1, z = a) = 0.625$, a researcher states:

"The data indicate that being a member of a Jazz band substantially increases the chance that a patient receiving the status quo drug lives 2 years after treatment. The estimated effect of being in a Jazz band is to increase the probability of living for 2 years after receiving the status quo drug from 0.25 to 0.625."

Does this statement accurately describe the empirical finding? Explain

Part (b): (30 Marks)

The *Kuznets filter* for annual data is given by:

$$c(L) = b(L)a(L)$$

where

$$\begin{aligned} a(L) &= \frac{1}{5}(L^{-2} + L^{-1} + L^{0} + L^{1} + L^{2}) \\ b(L) &= (L^{-5} - L^{5}) \end{aligned}$$

i. Given a frequency of 0.3, calculate the period p of the Kuznets filter. Show your calculations.

ii. Show that the gain function $|c(e^{-i\omega})|$ for the Kuznets filter can be expressed as

$$\frac{\sqrt{2}}{5}\sqrt{1 - i\sin(10\omega)}\sqrt{5 + 8\cos(\omega) + 6\cos(2\omega) + 4\cos(3\omega) + 2\cos(4\omega)}$$

3

Question 2 (100 Marks) -- Univariate Time Series & Forecasting.

Consider the moving average process of order q (MA(q))

$$y_t = \epsilon_t + \theta_1 \epsilon_{t-1} + \cdots + \theta_q \epsilon_{t-q} \quad t = 1, \ldots, T \quad \epsilon_t \sim iid(0, \sigma^2)$$

i. Derive the autocorrelation function for the MA(2) model.

Figure 1 shows the evolution of US real GDP.

ii. From visual inspection, does this series appear to be stationary or not and what might you expect if you plotted the correlogram for the series? Explain.

Now suppose someone suggested the following deterministic trend model to represent the time series of US real GDP:

$$y_t = \alpha + \beta t + u_t$$

iii. How could you induce stationarity in such a model and what might the process be called? Explain.

iv. If you were presented with the following autocorrelation functions and partial autocorrelation functions, what processes might you identify in figure 2 and figure 3? Explain. How might you estimate the process in figure 2? Explain. How might you estimate the process in figure 3? Explain.

v. Demonstrate the short memory or mean reversion of the MA(1) process:

$$y_t = \epsilon_t + \theta_1 \epsilon_{t-1}$$

by deriving the one- and two-period-ahead forecasts assuming invertibility (i.e. $|\theta_1| < 1$) and zero mean $(\mu = 0)$ assuming $\epsilon \sim iid(0, \sigma^2)$. What is the prediction mean square error for each of the two forecasts you made? Show your workings.

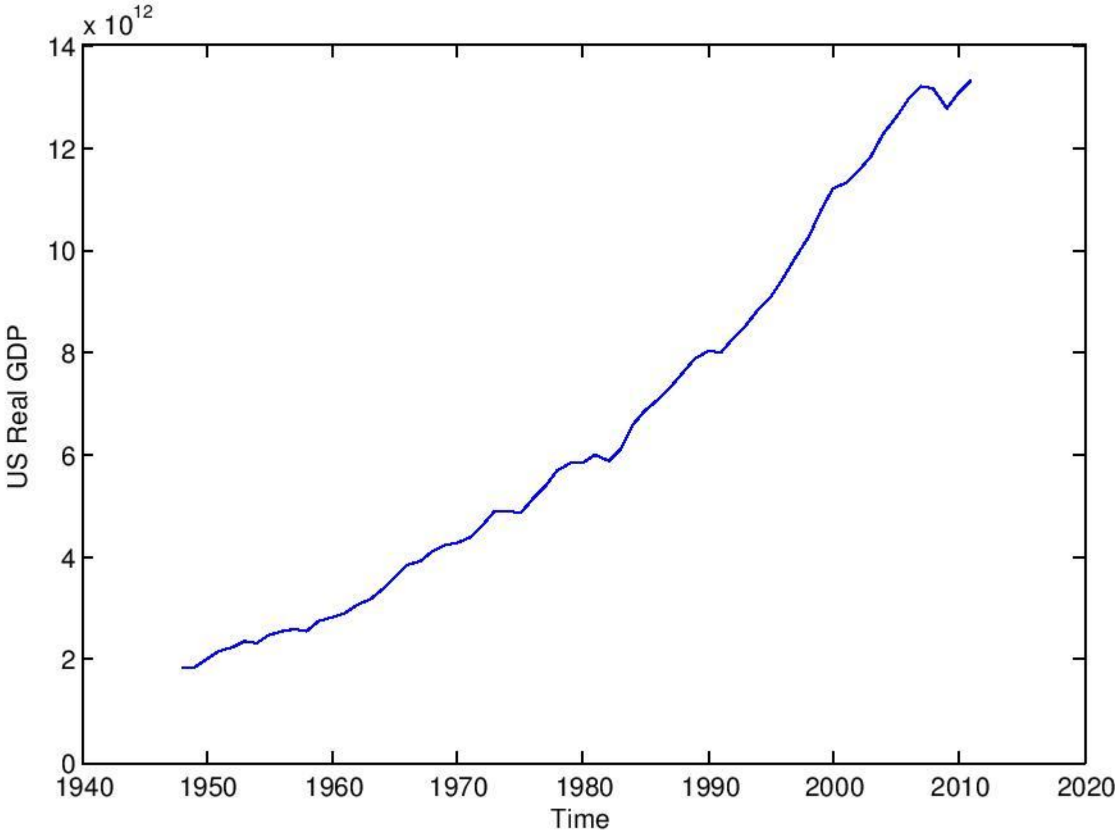

Figure 1: US Real GDP, 2005 US Dollars, Seasonally Adjusted. Source: IMF IFS.

5

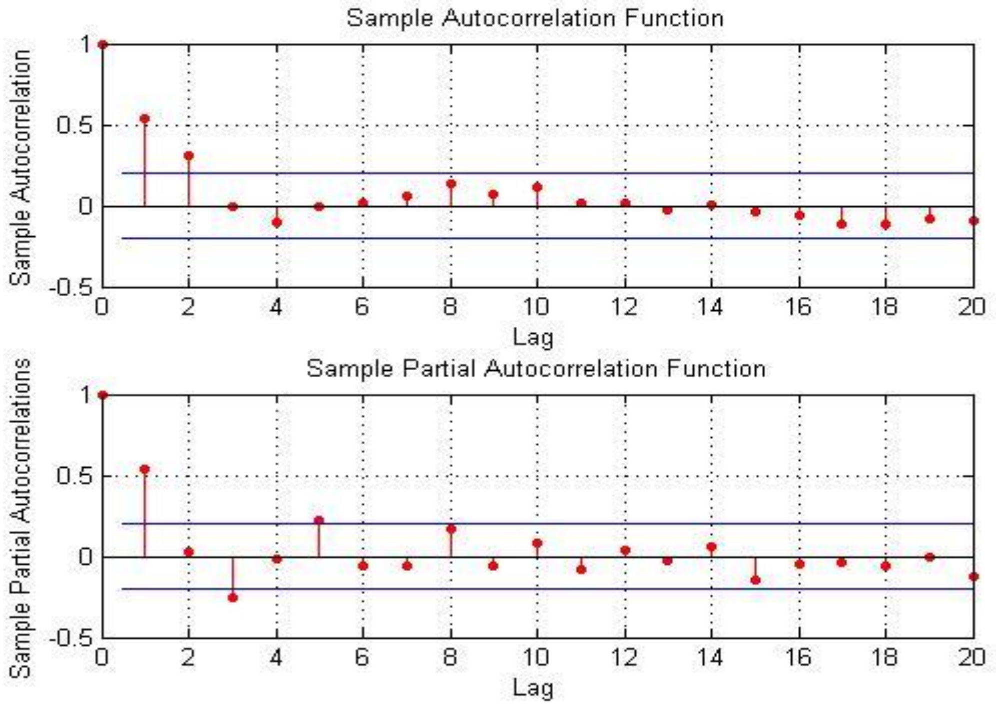

Figure 2: sample autocorrelation function and sample partial autocorrelation function for one set of time series data.

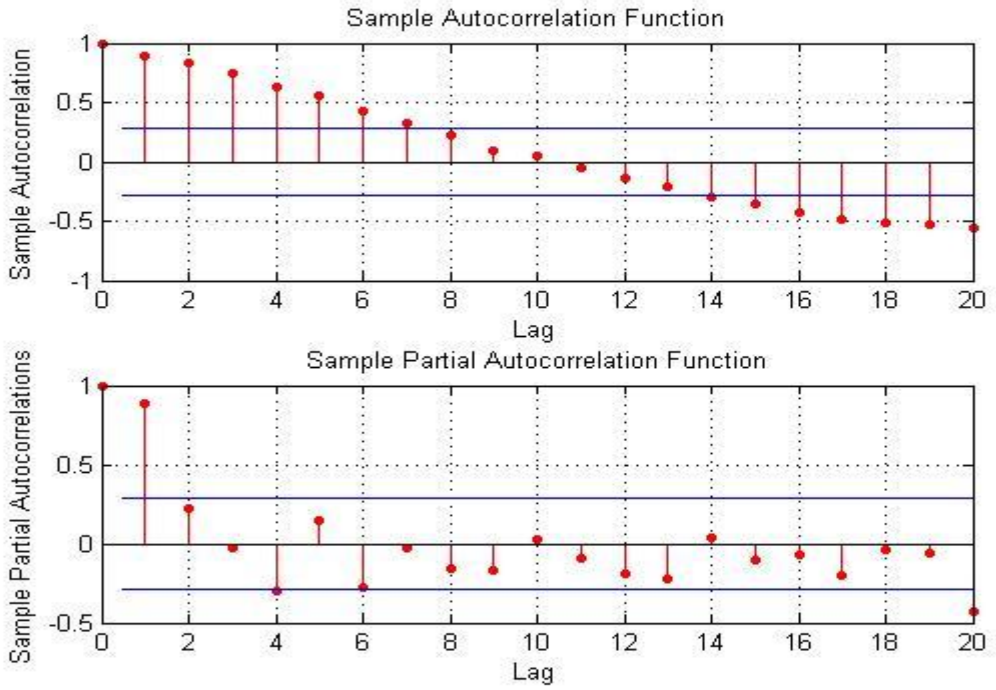

Figure 3: sample autocorrelation function and sample partial autocorrelation function for another set of time series data.

6

Question 3 (100 Marks) -- Volatility & Kalman Filtering.

Part (a): (20 Marks)

As discussed in lectures, most macro time series displays slowly evolving volatility. The stochastic volatility model is particularly apt for reflecting this stylised feature. Unlike GARCH, we model variance directly as an autoregressive stochastic process. This question asks you to demonstrate the fat tailed behaviour of the stochastic volatility model, which is a desirable property for many financial models. Recall the general stochastic volatility model from lectures

$$y_t = \sigma_t \epsilon_t \quad \sigma_t^2 = \exp(h_t) \quad t = 1, \ldots, T$$
$$h_t = \gamma + \phi h_{t-1} + \eta_t \quad \eta_t \sim NID(0, \sigma_\eta^2)$$

where we will assume that $\epsilon \sim N(0,1)$. Assume throughout that η_t is independent of ϵ_t.

Use the lemma that when $\exp(h_t)$ is log-normal, then the j^{th}moment around the origin is $\exp\{j\,\gamma_h + \frac{1}{2} j^2 \sigma_h^2\}$ to show that

$$V(y_t) = E(\epsilon_t^2) E\{\exp(h_t)\} \quad = \exp\{\gamma_h + \frac{1}{2}\sigma_h^2\}$$
$$E(y_t^4) = E(\epsilon_t^4) E\{\exp(2h_t)\} \quad = 3\exp\{2\gamma_h + 2\sigma_h^2\}$$

and from this show that when $\sigma_h^2 > 0$, that the model displays excess kurtosis as compared with the standard Normal distribution, where kurtosis is defined by the ratio

$$\frac{E[(Y – E(Y))^4]}{(E[(Y – E(Y))^2])^2}$$

Note that the standard Normal distribution has a kurtosis of 3.

Part (b): (80 Marks)

Recall from the definitions in lectures that the measurement equation and transition equations for the state space forms of models are given by

$$\mathbf{y}_t = \mathbf{Z}_t\boldsymbol{\alpha}_t + \mathbf{d}_t + \boldsymbol{\epsilon}_t \;\; t = 1, \ldots, T$$
$$\boldsymbol{\alpha}_t = \mathbf{T}_t\boldsymbol{\alpha}_{t-1} + \mathbf{c}_t + \mathbf{R}_t\boldsymbol{\eta}_t \;\; t = 1, \ldots, T$$

respectively, where $E(\boldsymbol{\alpha}_t) = \mathbf{a}_t$ where $\boldsymbol{\alpha}_t$ is the state vector and $Var(\boldsymbol{\alpha}_t) = \mathbf{P}_t$.

$$E(\boldsymbol{\epsilon}_t) = E(\boldsymbol{\eta}_t) = \mathbf{0} \quad Var(\boldsymbol{\epsilon}_t) = \mathbf{H}_t \quad Var(\boldsymbol{\eta}_t) = \mathbf{Q}_t$$

The Kalman filter prediction equations are given by

$$\mathbf{a}_{t|t-1} = \mathbf{T}_t\mathbf{a}_{t-1} + \mathbf{c}_t$$
$$\mathbf{P}_{t|t-1} = \mathbf{T}_t\mathbf{P}_{t-1}\mathbf{T}_t' + \mathbf{R}_t\mathbf{Q}_t\mathbf{R}_t' \;\; t = 1, \ldots, T$$

Note that the corresponding estimator of $\boldsymbol{y}_t$ is

$$\tilde{\mathbf{y}}_{t|t-1} = \mathbf{Z}_t\mathbf{a}_{t|t-1} + \mathbf{d}_t \;\; t = 1, \ldots, T$$

The prediction error or innovation vector is

$$\mathbf{v}_t = \mathbf{y}_t - \tilde{\mathbf{y}}_{t|t-1} = \mathbf{Z}_t(\boldsymbol{\alpha}_t - \mathbf{a}_{t|t-1}) + \boldsymbol{\epsilon}_t \;\; t = 1, \ldots, T$$

The MSE of the prediction error is

$$\mathbf{F}_t = \mathbf{Z}_t\mathbf{P}_{t|t-1}\mathbf{Z}_t' + \mathbf{H}_t$$

The Kalman filter updating equations are given by

$$\mathbf{a}_t = \mathbf{a}_{t|t-1} + \mathbf{P}_{t|t-1}\mathbf{Z}_t\mathbf{F}_t^{-1}(\mathbf{y}_t - \mathbf{Z}_t\mathbf{a}_{t|t-1} - \mathbf{d}_t)$$
$$\mathbf{P}_t = \mathbf{P}_{t|t-1} - \mathbf{P}_{t|t-1}\mathbf{Z}_t'\mathbf{F}_t^{-1}\mathbf{Z}_t\mathbf{P}_{t|t-1} \;\; t = 1, \ldots, T$$

Consider the following MA(1) model

$$y_t = \epsilon_t + \theta\epsilon_{t-1}$$

which can be written in state space form as

$$y_t = (1 \;\; 0)\boldsymbol{\alpha}_t \qquad\qquad t = 1, \dots, T$$

$$\boldsymbol{\alpha}_t = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \boldsymbol{\alpha}_{t-1} + \begin{bmatrix} 1 \\ \theta \end{bmatrix} \epsilon_t$$

where the state vector $\boldsymbol{\alpha}_t = (y_t, \theta\epsilon_t)'$. Note that if we let $\boldsymbol{\alpha}_t = (\alpha_{1t}, \alpha_{2t})'$, then $\alpha_{2t} = \theta\epsilon_t$ and $\alpha_{1t} = \alpha_{2,t-1} + \epsilon_t = \epsilon_t + \theta\epsilon_{t-1}$ and that the space state representation is not unique.

i. How would you initialise the Kalman filter? Explain.

ii. Write down the initial state vector $\mathbf{a}_0$, the initial MSE matrix $\mathbf{P}_0$, the first prediction error v_1 and its MSE f_1. Show your workings.

iii. What would the updated state vector $\mathbf{a}_1$ and covariance $\mathbf{P}_1$ be? Show your workings.

iv. Write the prediction equations $\mathbf{a}_{2|1}$ and $\mathbf{P}_{2|1}$ and v_2 and f_2. Show your workings.

9

Section B

Please attempt TWO questions from the three questions in this section.

Question 1 (100 marks) SUR and SEM models

Part **(a):** (25 marks)

Explain how the FGLS method is implemented in the context of SUR models. Under what circumstances would one estimate a system of equations by FGLS? What are the advantages of this approach? Why are VAR models estimated using equation-by-equation OLS as opposed to FGLS?

Part **(b):** (25 marks)

Consider the following Macro model

$$c_t = \alpha_0 + \alpha_1 y_t + \alpha_2 c_{t-2} + \varepsilon_{t1}$$

$$i_t = \beta_0 + \beta_1 r_t + \beta_2 (y_t - y_{t-1}) + \varepsilon_{t2}$$

$$y_t = c_t + i_t + g_t$$

where c_t is private consumption, i_t is investment, y_t is GDP, g_t is government spending and r_t is interest rate. Derive the completeness condition and discuss its implications for the identification of the above model.

Part **(c):** (25 marks)

Explain what is meant by observational equivalence. Discuss the main type of restrictions that are used to identify a simultaneous equation model.

Part **(d):** (25 marks)

Consider the following simultaneous equation model

$$y_1 = \gamma_1 y_2 + \beta_{11} x_1 + \beta_{21} x_2 + \beta_{31} x_3 + \varepsilon_1$$

$$y_2 = \gamma_2 y_1 + \beta_{12} x_1 + \beta_{22} x_2 + \beta_{32} x_3 + \varepsilon_2$$

Determine whether the following restrictions are sufficient to identify (or partially identify) the model: (i) $\beta_{21} = \beta_{32} = 0$, (ii) $\gamma_1 = 0$

Question 2 (100 marks) VAR models #1

Part **(a):** (30 marks)

Consider the following VAR model in structural form

$$y_t = b_{10} - b_{12} g_t + \gamma_{11} y_{t-1} + \gamma_{12} g_{t-1} + \varepsilon_{yt}$$

$$g_t = b_{20} - b_{21} y_t + \gamma_{21} y_{t-1} + \gamma_{22} g_{t-1} + \varepsilon_{gt}$$

where y_t is GDP and g_t is government consumption. Write the system in matrix form. Explain why this system cannot be estimated as is. Write the reduced form version of this model and write the reduced from errors as a function of the true structural shocks. Discuss how you would implement the Blanchard-Perotti (2002) identification approach in the context of this model. What are the underlying assumptions?

Part **(b)**: (30 marks)

Taking the model above, derive the vector moving average representation and discuss the assumptions that are needed to derive it.

Part **(c)**: (40 marks)

Using the Blanchard-Perotti (2002) identification assumptions derive the expression for the impact multiplier for a unit shock to government spending on GDP and for the effect a 1 unit shock to GDP on government spending.

Question 3 (100 marks) VAR models #2

Part **(a)**: (30 marks)

In the context of a VAR model with shocks identified via Choleski decomposition: explain what the variance decomposition is and how it is computed. Be as formal as possible.

Part **(b)**: (30 marks)

Discuss the different alternatives for computing the error bands for the impulse-response functions.

Part **(c)**: (40 marks)

Explain the main characteristics of the Choleski decomposition and the sign restriction approaches for the identification of shocks. Discuss their advantages and disadvantages.