University of Pennsylvania **Economics 706, Spring 2017**

Prelim Examination

Friday June 9, 2017. Time limit: 150 minutes

Instructions:

(i) The total number of points is 44, the number of points for each problem is given below.

(ii) The exam is closed book and closed notes.

(iii) To receive full credit for your answers you have to explain your calculations. **You may state additional assumptions.**

Problem 1: Two-Step-Ahead Forecasting (13 Points)

Consider the following data generating process:

$$y_t = \rho y_{t-1} + \epsilon_t + \theta\epsilon_{t-1}, \quad 0 \le \rho < 1, \quad \epsilon_t \sim iid(0,1). \tag{1}$$

A forecaster considers the following (misspecified) AR(1) model:

$$y_t = \phi y_{t-1} + u_t, \quad u_t \sim iid(0, \sigma_u^2). \tag{2}$$

(i) (2 Points) Suppose that the forecaster's model were correct and ϕ were known to the forecaster. Show that the optimal two-step ahead predictor under the quadratic forecast error loss function $L(e_{T+2|T}) = e^2_{T+2|T}$, where $e_{T+2|T} = y_{T+2} - \hat{y}_{T+2|T}$, is

$$\hat{y}^*_{T+2|T} = \phi^2 y_T. \tag{3}$$

(ii) (4 Points) Now suppose that the forecaster tries to estimate $\psi = \phi^2$ in (3) by direct estimation as follows. Let

$$\hat{\psi}_T = \text{argmin} \sum_{t=2}^{T} (y_t - \psi y_{t-2})^2. \tag{4}$$

Find the probability limit ψ_* of $\hat{\psi}_T$ under the assumption that $y_1, \ldots, y_T$ have been generated from model (1). Does this probability limit coincide with ρ^2 if $\theta = 0$?

(iii) (4 Points) Provide a characterization of the asymptotic variance of

$$\sqrt{T}(\hat{\psi}_T - \psi_*).$$

(iv) (3 Points) Approximate the frequentist prediction risk

$$\mathbb{E}_{\rho,\theta}[(y_{T+2} - \hat{\psi} y_T)^2]$$

by decomposing it into a part that is due to the conditional variance of y_{T+2}, a part that is due to the misspecification of the forecaster's model, and a part that is due to sampling variance of $\hat{\psi}_T$. Explain in what sense your calculation is approximate rather than exact.

Problem 2: Bayesian Estimation of AR(2) Model (11 Points)

Consider the process

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + u_t, \quad u_t \sim iidN(0,1) \tag{5}$$

and the prior distribution

$$\phi = [\phi_1, \phi_2]' \sim N\left(\mu, \frac{1}{\lambda} I\right), \tag{6}$$

where I is the identity matrix.

(i) (2 Points) For which values of ϕ is the process $\{y_t\}$ non-explosive?

(ii) (5 Points) Using a conditional likelihood function, derive the posterior distribution of $\phi|(Y_{1:T}, y_{-1}, y_0, \lambda)$, which we denote by $p(\phi|Y, \lambda)$. Interpret the formula for the posterior mean. What happens as you vary λ?

(iii) (4 Points) Derive the marginal likelihood (or data density) $p(Y|\lambda)$. How can $p(Y|\lambda)$ be used to determine the scaling of the prior distribution in a data-driven way?

Problem 3: Missing Observations (13 Points)

Consider the AR(1) model

$$x_t = \phi x_{t-1} + u_t, u_t \sim N(0,1). \tag{7}$$

Suppose that we observe $y_t = x_t$ if t is even ($t = 0, 2, 4, \dots, T$) and we don't have observations if t is odd ($t = 1, 3, 5, \dots, T-1$). This is a simplified version of a setting in which the period t is, say, a month, but we only observe the series, say, once a quarter.

(i) (2 Points) Taking y_{t-2} as given, derive $p(y_t|y_{t-2}, \phi)$.

(ii) (1 Points) Derive the (conditional) likelihood function

$$p(y_2, y_4, \dots, y_{T-2}, y_T|y_0, \phi).$$

(iii) (2 Points) Combine the likelihood function with a prior density of the form $\phi \sim N(0,1)$. Is the posterior distribution of ϕ normally distributed? Does this posterior appear to be more or less informative with respect to ϕ compared to a posterior based on observations for both even and odd time periods?

(iv) (2 Points) Explain why

$$p(x_{t-1}|y_t, y_{t-2}, \phi) = p(x_{t-1}|y_0, y_2, y_4, \dots, y_T, \phi). \tag{8}$$

(v) (4 Points) Derive $p(x_{t-1}|y_t, y_{t-2}, \phi)$ under the assumption that x_t is stationary and was initialized in the infinite past. Start by deriving the joint distribution $p(x_{t-1}, y_t, y_{t-2}|\phi)$ and then use the formula for the conditional mean $\mu_{x|y}$ and variance $\Sigma_{x|y}$ of a normal distribution.

(vi) (2 Points) Explain how one can use the ideas of data augmentation and Gibbs sampling to implement Bayesian inference in this model.

Problem 4: Vector Autoregressions (7 Points)

Consider a bivariate structural vector autoregression written in first differences:

$$\Delta y_t = \Phi \Delta y_{t-1} + \Phi_\epsilon \epsilon_t, \quad \epsilon_t \sim iidN(0, I). \tag{9}$$

Suppose the y_t is composed of log GDP and the log consumer price index (CPI). Moreover, suppose that we would like to interpret $\epsilon_{1,t}$ as monetary policy shock and $\epsilon_{2,t}$ as innovation to technology.

(i) (1 Points) Without further restrictions, are Φ and Φ_ϵ identifiable?

(ii) (2 Points) According to model (9) what is the effect of a unit shock $\epsilon_{1,t} = 1$ on Δy_{t+h}? Hint: you can introduce the vector $q = [1, 0]'$ to provide a formula and use the notation $\partial \Delta y_{t+h} / \partial \epsilon_{1,t}$.

(iii) (2 Points) What is the effect of a unit shock $\epsilon_{1,t} = 1$ on the level of y_{t+h}, that is, $\partial y_{t+h} / \partial \epsilon_{1,t}$?

(iv) (2 Points) What restriction is imposed on Φ_ϵ by the assumption that "monetary policy shocks do not have any effects on the level in the long-run?" Hint: you can interpret long-run as $h \longrightarrow \infty$. Does this assumption allow you to identify all the coefficients of the bivariate structural VAR in (9)?