

PRELIMINARY EXAM FOR THE Ph.D. DEGREE

Answer 4 questions, at least one from each part. Closed Book exam.

PART I

Question 1. Point allocation: (a) 10%; (b) 20%; (c) 20%; (d) 30%; (d) 20%.

If y takes only non-negative integer values and has geometric density with parameter λ then the density $f(y|\lambda)$ is

$$f(y) = \lambda^y(1 + \lambda)^{-(y+1)}, \quad y = 0, 1, \dots, \quad \lambda > 0.$$

Furthermore $E[y] = \mu$ and $V[y] = \mu(1 + \mu)$.

Here we introduce regressors and suppose that in the true model

$$E[y_i|\mathbf{x}_i] = \exp(\mathbf{x}_i'\boldsymbol{\beta}_0),$$

where $\boldsymbol{\beta}_0$ is an unknown $k \times 1$ parameter vector and \mathbf{x}_i is a $k \times 1$ nonstochastic regressor vector. We have a random sample (y_i, \mathbf{x}_i) , $i = 1, \dots, N$.

For much of this question we are concerned with the properties of the MLE of $\boldsymbol{\beta}$ under conditions weaker than correct specification of the density.

You can apply any laws of large numbers and central limit theorems without formally verifying the necessary assumptions.

(a) Give the formula for the objective function $Q_N(\boldsymbol{\beta})$ equal to N^{-1} times the log-likelihood function.

(b) Obtain $\text{plim } Q_N(\boldsymbol{\beta})$.

(c) Given your answer in (b), what assumptions are the essential assumptions to ensure consistency of $\hat{\boldsymbol{\beta}}$ that maximizes $Q_N(\boldsymbol{\beta})$.

(d) Assuming that the density is correctly specified, give the limit distribution of $\hat{\boldsymbol{\beta}}$. [Your derivation can be as brief as possible].

(e) Now suppose that the conditional density is misspecified, the associated conditional variance is misspecified, but the conditional mean is correctly specified. Give the formula for a consistent estimate of the asymptotic variance-covariance matrix of $\hat{\boldsymbol{\beta}}$. [Your answer and derivation can be as brief as possible].

Question 2. Point allocation: (a) 15%; (b) 15%; (c) 15%; (d) 15%; (e) 20%; (f) 10%; (g) 10%.

2. Consider a linear regression model based on the latent variable

$$y_i^* = \mathbf{x}_i' \boldsymbol{\beta} + u_i,$$

where the errors u_i are logistic distributed with density and cdf given by

$$f(u) = \frac{e^{-u}}{(1 + e^{-u})^2} \quad \text{and} \quad F(u) = \frac{1}{1 + e^{-u}}.$$

Note that the density is symmetric, so $f(u) = f(-u)$. For the logistic

$$E[u|u \leq c] = c + \frac{\ln(1 - F(c))}{F(c)}.$$

In this question the variable y_i^* is not completely observed.

(a) Suppose we observe only

$$\begin{aligned} y_i &= 1 & \text{if } y_i^* \geq 0 \\ y_i &= 0 & \text{if } y_i^* < 0 \end{aligned}$$

Give with justification the objective function for a consistent estimator of $\boldsymbol{\beta}$.

(b) Suppose we observe only

$$\begin{aligned} y_i &= 2 & \text{if } y_i^* \geq \alpha \\ y_i &= 1 & \text{if } 0 \leq y_i^* < \alpha \\ y_i &= 0 & \text{if } y_i^* < 0. \end{aligned}$$

State a method to consistently estimate $\boldsymbol{\beta}$ and α . For this part you may provide little detail.

(c) Suppose we observe only

$$y_i = y_i^* \quad \text{if } y_i^* \geq 0$$

Give with justification the objective function for the MLE of $\boldsymbol{\beta}$.

(d) In the same situation as in part (c), give an alternative consistent estimator for $\boldsymbol{\beta}$ that uses nonlinear least squares.

(e) Suppose that in part (a) one or more of the regressors \mathbf{x} are endogenous. Give details on a method to consistently estimate $\boldsymbol{\beta}$.

(f) Suppose that in part (a) the data are instead panel data with

$$y_{it}^* = \alpha_i + \mathbf{x}_{it}' \boldsymbol{\beta} + u_{it}, \quad i = 1, \dots, N, \quad t = 1, \dots, T.$$

and we observe $y_{it} = 1$ if $y_{it}^* > 0$ and $y_{it} = 0$ if $y_{it}^* \leq 0$.

Will this complicate estimation of $\boldsymbol{\beta}$ if α_i is iid $(0, \sigma_\alpha^2)$? A brief answer will do.

(g) Consider the same situation as (f) except α_i is a parameter to be estimated. Will this complicate estimation of $\boldsymbol{\beta}$? A brief answer will do.

PART II FOLLOWS ON NEXT PAGE

PART II

Question 3. Point Allocation: (a) 20%; (b) 60%; (c) 20%

Let $\{y_t\}$ be generated for $t = 1, \dots, T$ by the process

$$y_t = \mu + S_t, \text{ where } S_t = \sum_{j=1}^t v_j \text{ and } v_t \stackrel{iid}{\sim} N(0, \sigma^2); S_0 = 0$$

Consider estimating by least-squares the parameters of the model

$$y_t = \mu + \rho y_{t-1} + v_t$$

Define the scaling matrix

$$C_T = \begin{pmatrix} T^{1/2} & 0 \\ 0 & T^{3/2} \end{pmatrix}$$

(a) then show that

$$\begin{aligned} C_T \begin{pmatrix} \hat{\mu} - \mu \\ \hat{\rho} - 1 \end{pmatrix} &= \begin{pmatrix} 1 & T^{-2} \sum_{t=1}^T y_{t-1} \\ T^{-2} \sum_{t=1}^T y_{t-1} & T^{-3} \sum_{t=1}^T y_{t-1}^2 \end{pmatrix}^{-1} \begin{pmatrix} T^{-1/2} \sum_{t=1}^T v_t \\ T^{-3/2} \sum_{t=1}^T y_{t-1} v_t \end{pmatrix} \\ &= B_T^{-1} \begin{pmatrix} T^{-1/2} \sum_{t=1}^T v_t \\ T^{-3/2} \sum_{t=1}^T y_{t-1} v_t \end{pmatrix} \end{aligned}$$

(b) Given that

$$\begin{aligned} T^{-2} \sum_{t=1}^T S_t &\xrightarrow{p} 0; & T^{-3/2} \sum_{t=1}^T S_{t-1} v_t &\xrightarrow{p} 0 \\ T^{-5/2} \sum_{t=1}^T t S_t &\xrightarrow{L} W_1 & T^{-3} \sum_{t=1}^T S_t^2 &\xrightarrow{L} W_2 \\ T^{-3} \sum_{t=1}^T S_t^2 &\xrightarrow{p} 0 \end{aligned}$$

where W_1 and W_2 are non-degenerate distributions, show that:

$$p \lim_{T \rightarrow \infty} B_T = \begin{pmatrix} 1 & \frac{1}{2} \mu \\ \frac{1}{2} \mu & \frac{1}{3} \mu^2 \end{pmatrix} = B \text{ and } T^{-3/2} \sum_{t=1}^T y_{t-1} v_t \xrightarrow{L} N\left(0, \frac{1}{3} \sigma^2 \mu^2\right)$$

(c) Since the asymptotic covariance between $T^{-1/2} \sum_{t=1}^T v_t$ and $T^{-3/2} \sum_{t=1}^T y_{t-1} v_t$ is $\frac{1}{2} \sigma^2 \mu$ show that:

$$\begin{pmatrix} T^{-1/2} \sum_{t=1}^T v_t \\ T^{-3/2} \sum_{t=1}^T y_{t-1} v_t \end{pmatrix} \xrightarrow{L} N \begin{pmatrix} 0 \\ 0 \end{pmatrix}; \sigma^2 B \text{ and therefore}$$

$$C_T \begin{pmatrix} \hat{\mu} - \mu \\ \hat{\rho} - 1 \end{pmatrix} \xrightarrow{L} N \begin{pmatrix} 0 \\ 0 \end{pmatrix}; \sigma^2 B^{-1}$$

Carefully state any theorems and assumptions you make.

Hints:

$S_t = S_{t-1} + v_t$. Also

$$T^{-1} \sum_{t=1}^T \left(\frac{S_{t-1}}{\sqrt{T}} \right) \xrightarrow{L} \int_0^1 W(r) dr \text{ so } T^{-1} \sum_{t=1}^T \left(\frac{S_{t-1}}{\sqrt{T}} \right)^2 \xrightarrow{L} \int_0^1 W^2(r) dr$$

$$\lim_{T \rightarrow \infty} T^{-(n+1)} \sum_{t=1}^T t^n = (n+1)^{-1}$$

$$T^{-1/2} \sum_{t=1}^T v_t \xrightarrow{L} \int_0^1 dW(r) = W(1) \sim N(0,1)$$

$$T^{-3/2} \sum_{t=1}^T t v_t \xrightarrow{L} \int_0^1 r dW(r) \sim N(0, \frac{1}{3})$$

$$T^{-1} \sum_{t=1}^T S_{t-1} v_t \xrightarrow{L} \int_0^1 W(r) dW(r) \text{ so that } T^{-3/2} \sum_{t=1}^T S_{t-1} v_t \xrightarrow{p} 0$$

$$T^{-3/2} \sum_{t=1}^T y_{t-1} v_t = T^{-3/2} \sum_{t=1}^T \mu(t-1) v_t + T^{-3/2} \sum_{t=1}^T S_{t-1} v_t \xrightarrow{L} \mu \int_0^1 r dW(r)$$

Alternatively, by conventional methods,

$$\begin{pmatrix} T^{-1/2} \sum_{t=1}^T v_t \\ T^{-3/2} \sum_{t=1}^T t v_t \end{pmatrix} \xrightarrow{L} \begin{pmatrix} T^{-1/2} \sum_{t=1}^T v_t \\ \mu T^{-3/2} \sum_{t=1}^T t v_t \end{pmatrix} \xrightarrow{L} N \begin{pmatrix} 0 \\ 0 \end{pmatrix}; \sigma^2 B$$

Question 4. Point allocation: (a) 10%; (b) 5%; (c) 20%; (d) 15%; (e) 20%; (f) 30%

Consider the following stationary and ergodic data generating process for the random variable y_t :

$$y_t = \rho y_{t-1} + \varepsilon_t$$

where ε_t is a martingale difference sequence with $E(\varepsilon_t^2) = \sigma^2 > 0$ and $\sup_t E[|\varepsilon_t|^{\delta+2}] < \infty$ for $\delta > 0$; $|\rho| < 1$.

- (a) Obtain $E[y_t^2]$
- (b) Given that y_t is weakly stationary and ergodic, invoke an appropriate theorem to show that,

$$p \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T y_t^2 \rightarrow E[y_t^2]$$

- (c) Given (a) and (b) and knowledge that ε is a M.D.S., derive the limiting distribution of the least squares estimator for ρ . Please be precise in invoking all the relevant theorems that you require for your derivation.
- (d) Now suppose you are interested in the impulse response for period two, ρ^2 . Given the asymptotic distribution that you just derived for $\hat{\rho}_{LS}$, derive an asymptotic approximation for the distribution of ρ^2 .
- (e) Since the previous derivation is only a first order approximation to the true distribution, you decide to use the bootstrap to obtain a better approximation for the standard error of ρ^2 . Describe the steps involved in calculating an appropriate bootstrap estimator for this moment.
- (f) Instead, suppose you estimate ρ^2 with the projection of y_{t+1} on y_{t-1} . Derive the asymptotic distribution of the resulting least squares estimate and compare with your answer in (e). Briefly, Is there anything you can do to improve the efficiency of this estimator?

PART III FOLLOWS ON NEXT PAGE

Part III

Question 5: Point allocation: (a) 25%; (b) 25%; (c) 25%; (d) 25%.

Suppose that as part of your dissertation you are estimating a linear regression model by ordinary least squares. You are concerned about parameter stability in your model and perform the F-tests suggested by Bai and Perron (Econometrica, 1998). Your results are printed below:

```
*****
Output from the Bai and Perron testing procedure
*****
a) supF tests against a fixed number of breaks
-----
The supF test for 0 versus 1 breaks (scaled by q) is: 6.32
The supF test for 0 versus 2 breaks (scaled by q) is: 18.64
The supF test for 0 versus 3 breaks (scaled by q) is: 17.55
-----
The critical values at the 5% level are (for k=1 to 3):
12.89 11.60 10.46
-----
b) Dmax tests against an unknown number of breaks
-----
The UDmax test is: 18.64
(the critical value at the 5% level is: 13.27 )
*****
The WDmax test at the 5% level is: 20.06
(The critical value is: 14.19 )
*****
c) supF(I+1|I) tests using global optimizers under the null
-----
The supF( 2 | 1 ) test is : 28.43
The supF( 3 | 2 ) test is : 7.17
*****
The critical values of supF(i+1|i) at the 5% level are (for i=1 to 3 ) are:
12.89 14.50 15.42
*****
```

- (a) Draw conclusions about the number of breaks in your sample.
- (b) Suppose you conclude that there are some breaks in the parameters of your model and you want to incorporate them into your analysis. One of your options is to treat the breaks as deterministic as in the Bai/Perron approach. A second option is to model the breaks as random draws from some distribution using, for example, a Markov switching model. Outline the advantages and disadvantages of each of these approaches.

(question 5 continued on next page)

- (c) Suppose you conclude that there are some breaks in the parameters of your model and you want to incorporate them into your analysis. One of your options is to estimate a so-called time varying parameters model, i.e.,

$$y_t = x_t'(\beta + \varphi_t) + \varepsilon_t,$$

$$\varphi_t = F\varphi_{t-1} + u_t,$$

where $\varepsilon_t \sim iid N(0, \sigma^2)$, $u_t \sim iid N(0, \Sigma)$, and $E(u_t, \varepsilon_t) = 0$. Describe how you would use the EM algorithm to estimate the parameters F , β , σ^2 , and Σ . In general, what will be the asymptotic properties of your estimates?

- (d) This question is unrelated to parts (a)-(c). Suppose that you are trying to estimate $E(y|X)$ using kernel regression. Explain in words why this estimation problem is difficult when X is of high dimension and discuss whether the problems associated with high dimensional X are more or less severe when y is binary than when y is a continuous random variable.

Question 6: Point allocation: (a) 25%; (b) 25%; (c) 25%; (d) 25%.

Suppose that the solution of a representative agent's portfolio allocation problem is given by the following Euler equation

$$1 = E_t \left[\beta \frac{U'(C_{t+1})}{U'(C_t)} \frac{(P_{t+1} + D_{t+1})}{P_t} \right]$$

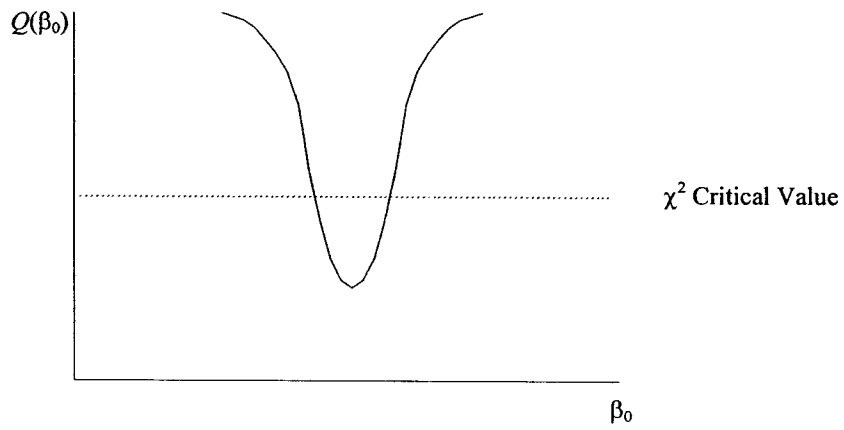
where C_t is consumption, P_t is the price of a financial asset, and D_t is the dividend paid out by the financial asset. Suppose that the agent's utility function is given by

$$U_t = C_t^{1/2}$$

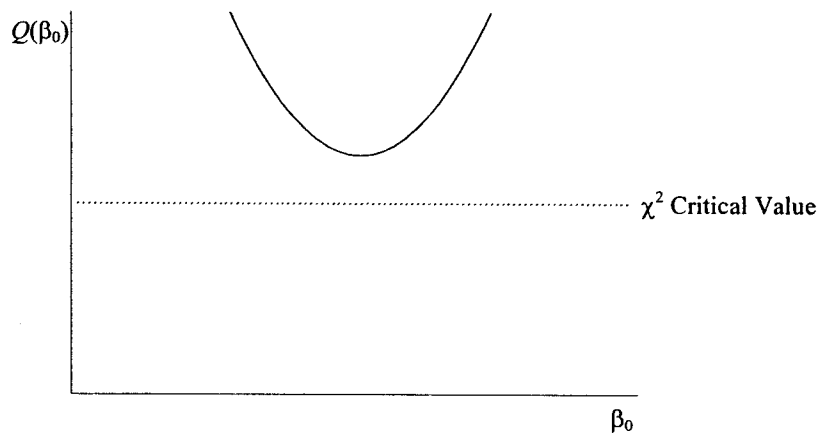
and you have quarterly time series data on P_t and D_t for one financial asset, and on C_t . In addition, you have data on the risk-free rate of return, R_t .

- (a) Write down the sample moment conditions that you would use to obtain the GMM estimates of the unknown parameter β . Be careful to justify your choice of moments.
- (b) Suppose that you estimate β using more than one moment condition. Given this set of moment conditions, explain how you would obtain the efficient GMM estimates.
- (c) Suppose that you were to estimate β using a least squares regression of 1 on X_{t+1} , where $X_{t+1} = (C_t / C_{t+1})^{1/2} (P_{t+1} + D_{t+1}) / P_t$. Show that this estimator is inconsistent and explain intuitively why it is inconsistent.
- (d) Let $Q(\beta_0)$ denote the efficient GMM objective function evaluated at the parameter value β_0 . Stock and Wright (Econometrica, 2000) showed that, under the null hypothesis that $\beta = \beta_0$, $Q(\beta_0)$ is asymptotically distributed as Chi-square with degrees of freedom equal to the number of moment conditions. They suggest that plots of $Q(\beta_0)$ against β_0 be used to reveal whether the parameter β is weakly identified. On the next page are three possible examples of such plots, labeled Example A, Example B, and Example C. For each of the three plots, draw conclusions about the properties of the model and the associated GMM estimates. Assume that the plots cover the entire domain of the parameter β .

Example A



Example B



Example C

