Fall 2007 R. Koenker

Economics 508

Midterm Exam Review

This is a merged and revised version of several recent midterm exams. The typical instructions are reproduced below. You can expect about 4 questions on the real midterm. You should bring a calculator in case there are questions that require some computation.

Please answer all questions. Even if you are unsure about some aspect of the questions, try to write something sensible – partial credit *will* be given. The questions will be weighted equally. The exam is closed book, closed notes and will last 2 hours.

1. You have just been hired by Starbuck's to forecast world coffee prices. Your predecessor has left a short paper describing his methods, which boil down to the estimated least squares equation

$$\log p_t = \underset{(0.23)}{-3.65} - \underset{(0.12)}{1.22} \log q_t + \underset{(0.42)}{2.86} \log x_t$$

where q_t, p_t and x_t are quantity, price and per capita income in billions of tons, dollars per pound, and 1000's of dollars per year, respectively. (Standard errors in parentheses.)

You feel that this is too simplistic, so you reestimate a more sophisticated model and obtain,

$$\log p_t = \underset{(0.18)}{-1.54} + \underset{(0.12)}{0.50} \log p_{t-1} + \underset{(0.07)}{0.28} \log p_{t-2} - \underset{(0.04)}{0.12} \log q_t - \underset{(0.02)}{0.05} (\log q_t)^2 + \underset{(0.31)}{1.04} \log x_t$$

It is presumed that world coffee production is determined primarily by weather conditions in growing areas, so it is reasonable to treat supply as exogonous and view the foregoing models as proper demand equations, but this year there has been talk about the possibility that the Organization of Coffee Exporting Nations will conspire to reduce world output by 5% from its current level of 2.718 trillion tons.

   (a) Evaluate the stability of the dynamics of your estimated model.

   (b) Compare the price predictions from the two models (from a long-run equilibrium viewpoint) based on the conjectured 5% reduction of supply.

2. In his canonical work on cointegration in multivariate (VAR) settings Soren Johansen focused on the error correction model

$$\Delta y_t = \Pi y_{t-1} + B_1 \Delta y_{t-1} + \cdots + B_p \Delta y_{t-p} + \varepsilon_t \tag{1}$$

where y_t is an m-vector.

(a) Suppose, for a moment, that $m = 1$, so that Π and the B's are all scalar, interpret the hypothesis $\Pi = 0$ and explain *briefly* how to test it.

(b) Now in the general case $m \geq 2$ explain why Johansen's test relies on the estimated auxiliary VAR equations:

$$(2a) \qquad \Delta y_t = \hat{\delta}_0 + \hat{\delta}_1 \Delta y_{t-1} + \cdots + \hat{\delta}_p \Delta y_{t-p} + \hat{u}_t$$

$$(2b) \qquad y_{t-1} = \hat{\theta}_0 + \hat{\theta}_1 \Delta y_{t-1} + \cdots + \hat{\theta}_p \Delta y_{t-p} + \hat{v}_t$$

In particular how does estimating these equations relate to the estimation of (1)?

(c) Give a brief heuristic explanation of what it means that the coordinates of y_t are *cointegrated* and how this relates to the rank of the matrix Π. (Avoid discussion of any gory details of the Johansen test statistic.)

3. Suppose that we have a structural equation

$$y = Y\gamma + X\beta + u$$

and we are willing to assume that $X \perp u$ but *not* $Y \perp u$. But we have instrumental variables $Z \perp u$ that might save the day. We proceed in the following somewhat unorthodox way:

Step 1 We multiply by $M_X = I - X(X'X)^{-1}X'$ to obtain

$$M_X y = M_X Y\gamma + M_X X\beta + M_X u$$

which we write as

$$\tilde{y} = \tilde{Y}\gamma + \tilde{u}$$

Step 2 Now, even though we know that Z doesn't belong in our structural equation and that Y does belong, but is endogenous, we consider estimating the linear model by least squares

$$\tilde{y} = \tilde{Y}\gamma + Z\delta + v$$

in particular, we consider a family of estimaters of δ as a function of γ,

$$\hat{\delta}(\gamma) = (Z'Z)^{-1}Z'(\tilde{y} - \tilde{Y}\gamma)$$

Step 3 Now since we are really interested (mainly) in γ we decide to estimate γ by minimizing the norm of $\hat{\delta}(\gamma)$ that is by solving

$$\hat{\gamma}_A = \arg\min_{\gamma} \parallel \hat{\delta}(\gamma) \parallel_A^2$$

where $\parallel X \parallel_A = x'Ax$.

(a) Explain Step 1. What happened to β in the model for $\tilde{y}$?

(b) Given that you are free to choose A in any way you would like, try to provide a justification for Steps 2 and 3. Be sure to be explicit about any reservations you have about this procedure, and suggest alternatives that would perform better if you can.

4. Suppose that

$$y = X_1\beta_1 + X_2\beta_2 + u$$

with $Euu' = \sigma^2 I$, and let

$$\tilde{\beta}_1 = (X_1'X_1)^{-1}X_1'y$$

and

$$\hat{\beta}_1 = (X_1'M_2X_1)^{-1}X_1'M_2y$$

(a) Explain *in words* the difference between these two estimators.

(b) Define the matrix M_2.

(c) It is claimed that

$$\hat{\beta}_1 = \tilde{\beta}_1 - \Gamma\hat{\beta}_2.$$

Derive this expression providing in the process definitions of $\hat{\beta}_2$ and Γ.

(d) It is claimed that

$$V(\hat{\beta}_1) = V(\tilde{\beta}_1) + \sigma^2\Gamma(X_2'M_1X_2)^{-1}\Gamma'.$$

Interpret this expression and explain briefly "what happened to the covariance term?"

(e) The foregoing results establish a fundamental tradeoff in the consideration of omitted variables between bias and variance effects. Explain briefly the nature of this tradeoff and contrast it briefly with the model selection rules proposed by Akaike and Schwarz in the case that β_2 is a scalar.

5. In Figures 2a-f we illustrate six scatterplots corresponding to possible models represented algebraically below. Try to match the figures and the equations; if there isn't a match try to explain why you don't think so.

$$\log y_i = \underset{(0.04)}{-2.00} + \underset{(0.026)}{2.01} \log x_i \tag{2.1}$$

$$y_i = \underset{(.056)}{5.03} - \underset{(.0.10)}{2.03} /\sqrt{x_i} \tag{2.2}$$

$$y_i = \underset{(1.03)}{49.92} - \underset{(0.13)}{1.91} x_i \tag{2.3}$$

$$1/y_i = \underset{(.030)}{1.047} - \underset{(.004)}{0.051} x_i \tag{2.4}$$

$$y_i = \underset{(.51)}{49.5} + \underset{(.065)}{1.95} x_i \tag{2.5}$$

$$y_i = \underset{(.013)}{2.05} - \underset{(.004)}{.060} x_i + \underset{(.0002)}{.004} x_i^2 \tag{2.6}$$

6. Suppose you have estimated a cost function for a sample of international steel manufacturers

$$\log c_i = \alpha + \beta \log x_i + \gamma(\log x_i)^2 + u_i$$

where c_i denotes total annual cost of firm i and x_i denotes annual output in metric tons.

(a) Explain why optimal scale (minimum average cost), if it exists, occurs at the output level where the cost elasticity is unity, i.e., $\eta = \partial \log c / \partial \log x = 1$,

(b) Having estimated this model explain how you would make a point estimate of optimal scale. State clearly any necessary caveats.

7. The Schwarz information criterion for model selection chooses the model that maximizes

$$\mathrm{SIC}_j = \ell_j(\hat{\theta}) - \frac{1}{2} p_j \log n$$

(a) Explain briefly what each of the pieces mean: $\ell_j(\hat{\theta}), p_j, n$.

(b) Explain why such a criterion might be preferred to finding the model that maximized the likelihood function as a device for model selection.

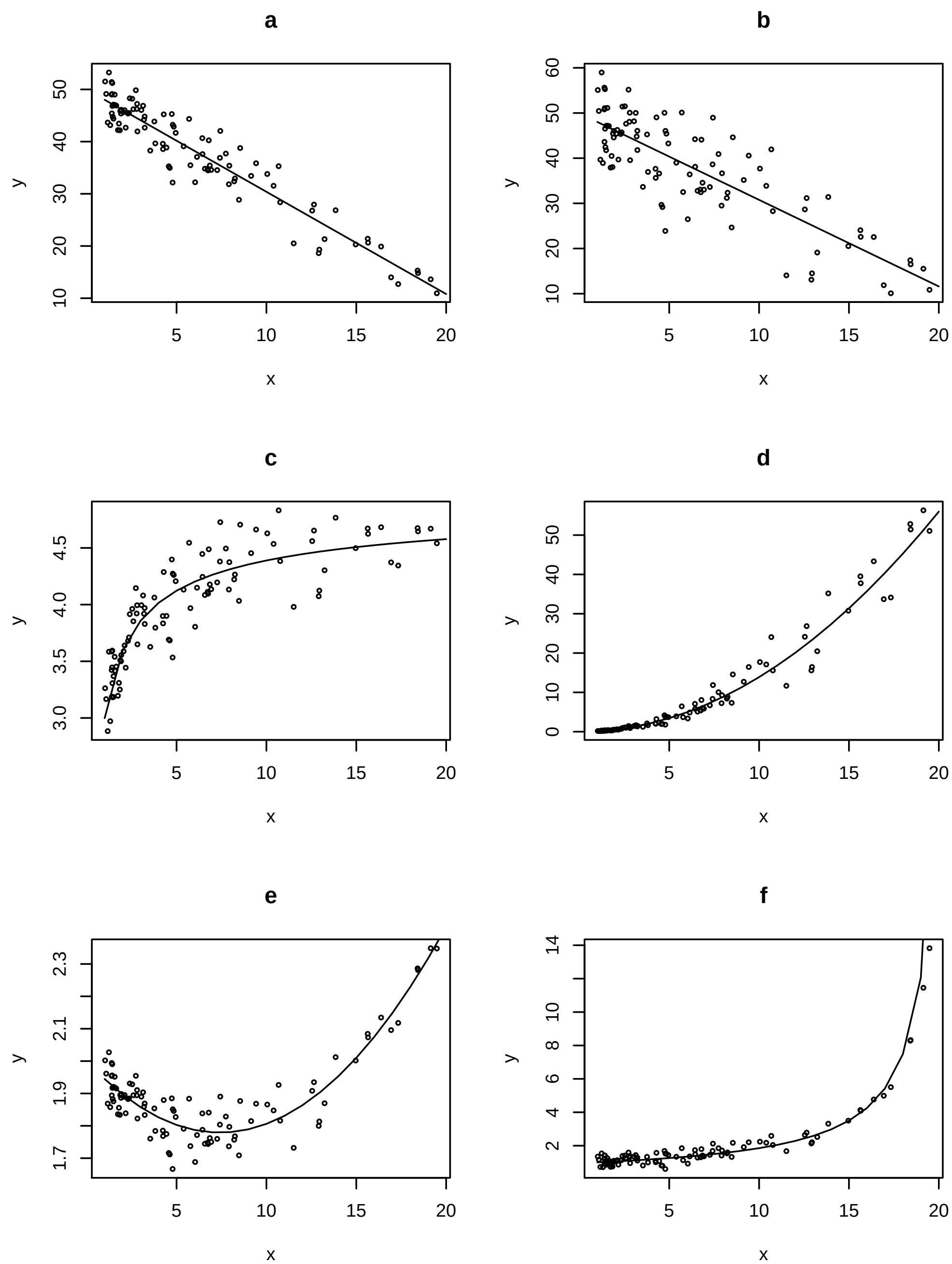

Figure 1: Six scatterplots and their fitted relationship.

(c) Explain how the SIC criterion is related to the classical theory of hypothesis testing. (Recall that under the null hypothesis that model j is correct, then for any model k within which model j is nested, i.e., $\Theta_k \supseteq \Theta_j$, we have

$$2(\ell_k(\hat{\theta}) - \ell_j(\hat{\theta})) \sim \chi^2_q$$

where $q = p_k - p_j$. In particular, relate SIC to the conventional t-test when $q = 1$, so the model dimensions only differ by one.

8. Consider the problem

$$(*) \qquad \min_{b \in \Re^p} (y - Xb)' A (y - Xb)$$

(a) Solve to obtain an explicit expression for the optimal value of the vector b.

(b) Formulate the following problems as special cases of problem (*), explicitly defining appropriate notation for each: (i) Generalized least squares, (ii) two stage least squares (iii) the partial residual plot.

9. Given the simple cobweb model of supply and demand

$$\begin{array}{ll} \text{(Supply)} & Q_t = \alpha_1 + \alpha_2 p_{t-1} + \alpha_3 z_t + u_t \\ \text{(Demand)} & p_t = \beta_1 + \beta_2 Q_1 + \beta_3 w_t + v_t \end{array}$$

Suppose that z_t and w_t are exogenous in the sense that $E(z_t, w_t)'(u_t, v_t) = 0$.

(a) Supposing all the variables are in logs, find the long-run supply elasticities with respect to changes in z_t and w_t.

(b) Explain briefly how to use the δ method to compute a standard error for the elasticities found in part a.)

10. A standard formula for the least squares estimator of scalar parameter β in the simple bivariate regression model

$$y_i = \alpha + \beta x_i + u_i$$

is

$$\hat{\beta} = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sum(x_i - \bar{x})^2}$$

Explain this formula using the language and algebra of the partial residual plot.

11. The Vichy bottled water company has estimated the following log quadratic demand curve

$$\log q_t = \hat{\alpha} + \hat{\beta} \log p_t + \hat{\gamma}(\log p_t)^2 + \hat{\rho} \log q_{t-1}$$

(a) Since costs are negligible, the firm wants to maximize *revenue*. Explain what you would do to find the revenue maximizing price, explaining also any caveats along the way.

(b) Not satisfied with a point estimate for this revenue maximizing price, p^*, suppose the Board of Directors wants a 95% confidence interval for p^*. Describe how you would compute the confidence interval, using the δ-method?

12. Consider a model of two U.S. interest rates: r_t is the rate on 90 day US Treasury Bills, and R_t the rate on one-year Treasury Bonds. Using quarterly data from 1962-99 one obtains augmented Dickey-Fuller test statistics including an intercept of -2.96 and -2.22 respectively, for the hypotheses that r_t and R_t have a unit root. In contrast, the difference between the two rates $(R_t - r_t)$ has an ADF statistic of -6.31. Interpret these results and the associated vector autoregressive model (standard errors appear in parentheses):

$$\begin{aligned}
\Delta r_t = \; & \underset{(0.17)}{0.14} - \underset{(0.32)}{0.24} \; \Delta r_{t-1} - \underset{(0.34)}{0.44} \; \Delta r_{t-2} \\
& - \underset{(0.09)}{0.01} \; \Delta R_{t-1} + \underset{(0.27)}{0.15} \; \Delta R_{t-2} \\
& + \underset{(0.07)}{0.18} \; (R_{t-1} - r_{t-1}) \\
\Delta R_t = \; & \underset{(0.16)}{0.36} - \underset{(0.30)}{0.14} \; \Delta r_{t-1} - \underset{(0.29)}{0.33} \; \Delta r_{t-2} \\
& - \underset{(0.35)}{0.11} \; \Delta R_{t-1} + \underset{(0.25)}{0.10} \; \Delta R_{t-2} - \underset{(0.14)}{0.52} \; (R_{t-1} - r_{t-})
\end{aligned}$$

Viewing the VAR model as a forecasting model, try to provide some sort of economic interpretation for it.

13. For an evaluation of mutual funds you have been asked to estimate models of quarterly rates of return for n firms over T quarters. You first estimate a model that looks like,

$$y_{i,t} = \mu_i + \alpha y_{i,t-1} + \beta_i x_t + u_{i,t} \tag{1}$$

where $y_{i,t}$ denotes the rate of return of fund i in quarter t and x_t denotes the market rate of return in quarter t.

(a) As a preliminary step you decide to test model (1) against the simpler model,

$$y_{i,t} = \mu + \alpha y_{i,t-1} + \beta x_t + v_{i,t} \tag{2}$$

Explain briefly how you would formulate the test given error sums of squares for the two models.

(b) Suppose that the null hypothesis fails to be rejected in part a.) so you decide to proceed with the simpler model and you obtain,

$$\begin{array}{rcccccl} \hat{y}_{i,t} = & 0.0014 & + & 0.862 & y_{i,t-1} + & 0.1481 & x_t \\ & (0.012) & & (0.021) & & (0.061) & \end{array} \tag{3}$$

with $\hat{\sigma}_v^2 = 0.00010$. If the market rate of return were stable at 0.05, what would you expect *based on the model* to be the unconditional mean and variance of the mutual funds returns at any given time t.

(c) Suppose you were to group funds using your mean and standard deviation from part b.) into performance groups at some initial time t_0, as illustrated in Table 1. You then follow the performance of the group over the next several years. What would you expect to see happen to mean performance of each group? What would happen to the variance of each group?

Grade	Rate of Return at t_0
Poor	$(-\infty, \hat{\mu} - 0.84\hat{\sigma}]$
Weak	$(\hat{\mu} - 0.84\hat{\sigma}, \hat{\mu} - 0.24\hat{\sigma}]$
Average	$(\hat{\mu} - 0.24\hat{\sigma}, \hat{\mu} + 0.24\hat{\sigma}]$
Good	$(\hat{\mu} + 0.24\hat{\sigma}, \hat{\mu} + 0.84\hat{\sigma}]$
Great	$(\hat{\mu} + 0.84\hat{\sigma}, \infty]$

Table 1: Mutual Fund Rating Scheme