Fall 2006 Professor 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 been asked to comment on a paper on the world demand for coffee at the summer meetings of the Juan Valdez Foundation. The estimated model looks like

$$\log p_t = \underset{(0.23)}{-3.67} - \underset{(0.12)}{1.42} \log q_t + \underset{(0.42)}{2.86} \log x_t$$

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

You have decided to focus your remarks on the implications of the model for recent proposals suggesting that coffee growers could increase their revenue if they could agree to destroy 5 or 10 percent of the annual crop, currently approximately 2.718 billion tons.

Based on models for similar commodities you are quite convinced that the static model reported above is misspecified. You reestimate the model and obtain:

$$\log p_t = -\underset{(0.18)}{1.54} + \underset{(0.12)}{0.78} \log p_{t-1} - \underset{(0.06)}{0.12} \log q_t - \underset{(0.02)}{0.05} (\log q_t)^2 + \underset{(.31)}{1.04} \log x_t$$

   (a) Compare the implications of your estimated model with those of the first model stressing their implications for the policy question raised above. In particular, compare the long-run equilibrium predictions of the two models, and contrast their equilibrium implications for the contemplated policy initiative.

   (b) Explain briefly how you would predict the equilibrium effect of a 10 percent crop reduction on aggregate revenue, and how you would go about estimating a standard error for this prediction.

(c) Suppose you wanted to test for autocorrelation of the errors in the revised demand equation how would you go about it?

(d) Now suppose world coffee supply is determined by an equation of the form

$$\log q_t = \alpha_0 + \alpha_1 \log q_{t-1} + \alpha_2 \log p_t + \alpha_3 \log z_t + \alpha_4 \log w_t + u_t$$

where z_t and w_t can be considered exogenous to the world coffee market. Explain briefly how this supply equation would alter your estimation strategy for the revised model.

2. 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.

3. 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.

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

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

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

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

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

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

4. 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.

5. 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.

(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

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

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.

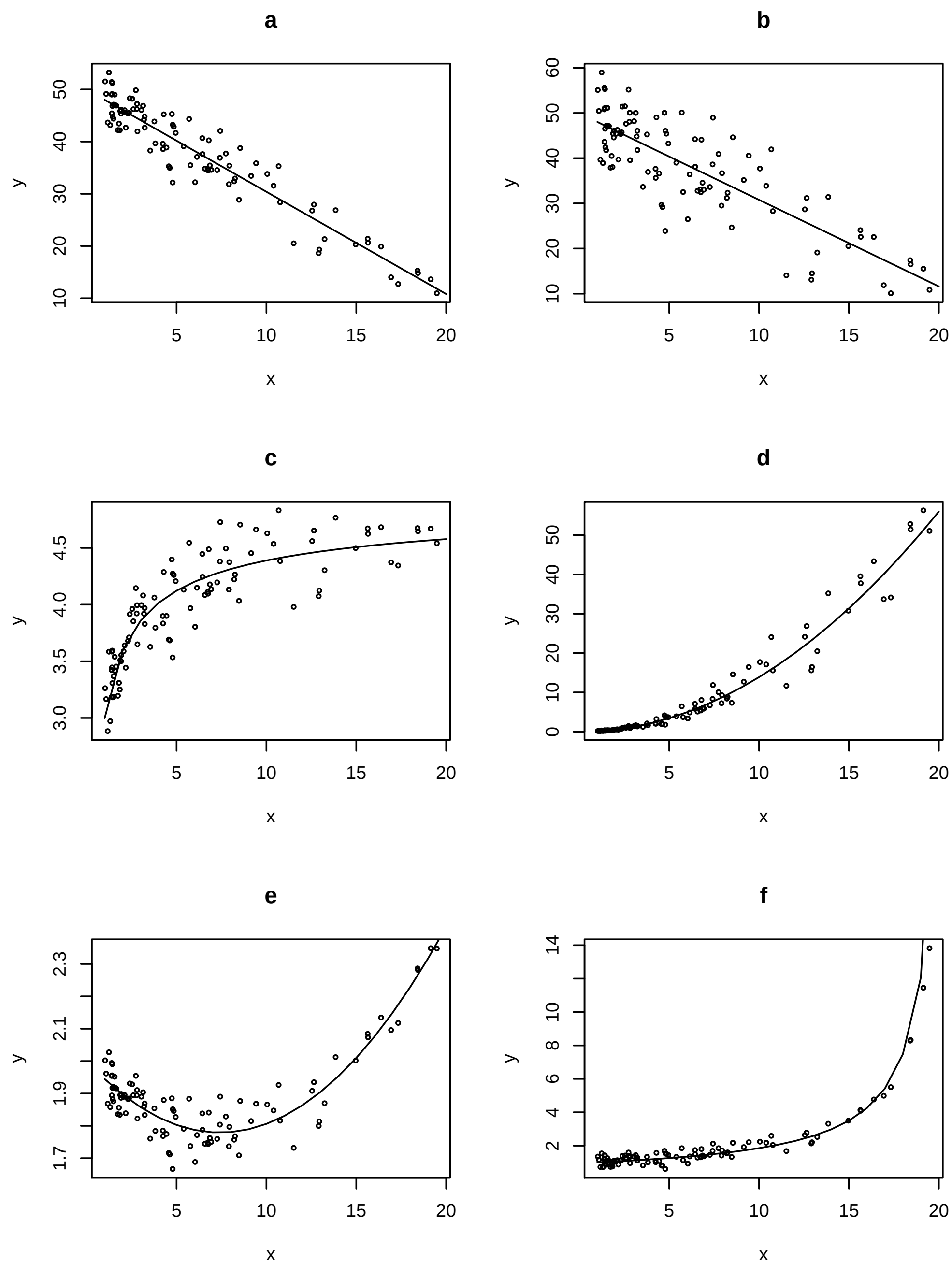

Figure 1: Six scatterplots and their fitted relationship.

6. 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.

7. 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.)

8. 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.

9. 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?

10. 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):

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

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

11. Suppose that we have the structural equation

$$y_1 \quad = \quad Y_1\gamma_1 + X_1\beta_1 + u_1 \equiv Z_1\delta_1 + u_1$$

and exogenous variables $X = [X_1 \vdots X_2]$. To estimate the parameter vector δ_1 we consider three alternative estimators:

$$\begin{aligned}
\hat{\delta}_{IV} &= (\hat{Z}_1' Z_1)^{-1} \hat{Z}_1' y_1 \\
\hat{\delta}_{2SLS} &= (\hat{Z}_1' \hat{Z}_1)^{-1} \hat{Z}_1' y_1 \\
\hat{\delta}_{CV} &= (Z_1' M_R Z_1)^{-1} Z_1' M_R y_1
\end{aligned}$$

where

$$\begin{aligned}
\hat{Z}_1 &= P_X Z_1 \\
P_X &= X(X'X)^{-1}X' \\
R &= Y_1 - \hat{Y}_1 = (I - P_X) Y_1 \\
M_R &= I - R(R'R)^{-1}R'
\end{aligned}$$

Explain briefly the rationale for $\hat{\delta}_{IV}$, and what assumptions are required to justify it. Then explain why $\hat{\delta}_{2SLS}$ and $\hat{\delta}_{CV}$ are equivalent to it. Conclude by explaining how to compute $\hat{\delta}_{CV}$ using a sequence of ordinary least squares steps.