# Review Session Problem Sets

ECON 343

Jad Chaaban

# Practice exercise 1

1. Using data on housing price and house characteristics we estimate the model,

$$price_i = \beta_0 + \beta_1 lotsize_i + \beta_2 sqrft_1 + \beta_3 bdrms_i + \beta_4 colonial_i + u_i, \quad (1)$$

where $price$ is sale price of house $i$ in \$1,000s, $lotsize$ is the size of lot in square feet, $sqrft$ is the size of the house in square feet, $bdrms$ is the number of bedrooms and $colonial$ is a dummy variable indicating that the house is Colonial style. The estimated coefficients are given in Column 1 of Table 1.

(a) Based on the reported estimates, answer the following questions

i. How much are people willing to pay for an additional bedroom?

ii. How much more or less are homebuyers willing to pay for a colonial style house?

iii. What unobservables might be included in the error term?

Table 1: Estimates for Model 1

| | (1) | (2) | (3) |
|---|---|---|---|
| Size of lot ($ft^2$) | 0.002 | 0.002 | 0.215 |
| | (0.001)** | (0.001) | (0.074)** |
| Size of house ($ft^2$) | 0.124 | 0.124 | 1.277 |
| | (0.013)** | (0.018)** | (1.531) |
| Number of Bedrooms | 11.004 | 11.004 | 1,637.589 |
| | (9.515) | (9.259) | (1,092.470) |
| Colonial (0/1) | 13.716 | 13.716 | -2,848.362 |
| | (14.637) | (16.430) | (1,680.540) |
| Constant | -24.127 | -24.127 | -4,995.030 |
| | (29.603) | (37.778) | (3,398.845) |
| Observations | 88 | 88 | 88 |
| R-squared | 0.68 | 0.68 | 0.18 |
| F(4, 83) | 43.25 | 18.24 | 4.66 |
| Prob>F | 0.0000 | 0.0000 | 0.0019 |

Standard errors in parentheses. * significant at 5%; ** significant at 1%

Dependent variable in columns 1 and 2 is price. Dependent variable in column 3 is $\hat{u}^2$

# Practice exercise 2

The following table contains the following information for 15 assets: beta coefficient (*beta*); standard measure of unsystematic risk ($\hat{\sigma}$); mean annual return ($\bar{R}$).

| Asset | *beta* | $\hat{\sigma}$ | $\bar{R}$ |
|---|---|---|---|
| Treasury Gilts (undated; 2½ %) | 0 | 0 | ? |
| Anglian Water Group | .24 | .18 | .08 |
| Scottish Power | .56 | .16 | .09 |
| Imperial Tobacco | .70 | .36 | .06 |
| Tesco | .74 | .07 | .06 |
| Zurich Finance | .92 | .17 | .09 |
| Prudential | .94 | .08 | .10 |
| Royal Bank of Scotland | 1.13 | .18 | .14 |
| Northern Rock | 1.18 | .15 | .13 |
| Cathay Pacific | 1.22 | .22 | .25 |
| Gillette | 1.39 | .14 | .12 |
| Newcastle United | 1.43 | .28 | .08 |
| Hilton | 1.64 | .37 | .37 |
| EuroDisney | 1.82 | .19 | .18 |
| LA Fitness | 2.44 | .18 | .14 |

(a) The price today of £100 nominal of Treasury Gilts (undated; 2½ %) is £52.20. Using this information, deduce the missing number in the table. [5]

(b) Explain why Treasury Gilts have both a zero beta and a zero $\hat{\sigma}$ . [4]

(c) Comment briefly on two other beta coefficients in the table. Are their values roughly what you would have expected for the stocks in question? [6]

(d) The data in the table were used to estimate the following model:

$$\hat{\bar{R}} = 0.015 + 0.054\,beta + 0.308\hat{\sigma}$$

$$(se)\ (0.045)(0.033) \qquad (0.207) \qquad\qquad (1)$$

$$n = 15$$

Use the results of this model to conduct a test of the CAPM. What does your test result imply about investor behaviour? [6]

# Practice exercise 3

Daily data on the price of shares in BSkyB were used to estimate the following models. $X_t$ is the share price on day $t$. The other variables are day-of-week dummies.

Model 1:

$$\hat{X}_t = 0.05 + 0.97X_{t-1}$$
$$(se)\ (0.03)\ (0.03)$$
$$n = 344 \quad R^2 = 0.87 \quad DW = 1.26$$

Model 2:

$$\hat{X}_t = 0.02 + 0.98X_{t-1} + 0.15Tues + 0.18Weds + 0.21Thurs + 0.24Fri$$
$$(se)\quad (0.04)\ (0.04)\qquad (0.08)\qquad (0.08)\qquad (0.08)\qquad (0.08)$$
$$n = 344 \quad R^2 = 0.89 \quad DW = 1.91$$

(a) Test for serial correlation in Model 1. Is your result consistent with the Efficient Markets Hypothesis? [6]

(b) Explain why one of the five day-of-week dummies has necessarily been excluded from Model 2. What would happen if all five were included? [6]

(c) By testing model 1 as a restricted version of model 2, test the significance of day-of-week in the determination of the price of BSkyB shares. [7]

(d) What is the weekly pattern in the share price? How would you exploit knowledge of this pattern in the construction of a profitable trading rule for BSkyB shares? [7]

(e) Test for serial correlation in model 2. Explain why serial correlation is *a priori* less likely to be apparent in Model 2 than in Model 1. [7]

# Practice exercise 4

Column 2 of Table 4 below reports least squares estimates of the model,

$$log(avgprc) = \beta_0 + \theta_0 t + \delta_0 Mon_t + \delta_1 Tues_t + \delta_2 Wed_t + \delta_3 Thurs_t + e_t, \qquad (3)$$

where $log(avgprc)$ is average daily price of fish at the Fulton Fish Market in Manhattan. The log of average price of fish is regressed a function of a linear time trend, $t$, and daily dummy variables, $Mon$ through $Thurs$. The error term is given by $e_t$.

(a) Why is a dummy variable for Friday excluded?

(b) Based on the estimates in Column 2 of Table 4, is there evidence that price varies systematically within a week? Test the daily dummies individually, and using the $R^2$ form of the F-test, test the joint significance of the daily dummies.

(c) In Column 3 of Table 4, we add the variables $wave2$ and $wave3$, measures of wave heights over the past several days, to test how rough seas affect fish prices. Are these variables individually significant? Describe how stormier seas would affect the price of fish. Are they jointly significant?

(d) What happened to the time trend when $wave2$ and $wave3$ were added to the regression? What must be going on?

(e) Explain why all the explanatory variables in the regression are safely assumed to be strictly exogenous.

(f) Table 5 reports the estimates of the residuals from the model in column 3 on a lag. What do you conclude about the presence of AR(1) serial correlation from these estimates?

Table 4: Least Squares estimates of Price of Fish

| | (1) | (2) | (3) |
|---|---|---|---|
| Time trend | -0.004 | -0.004 | -0.001 |
| | (0.001)** | (0.001)** | (0.001) |
| Monday | | -0.010 | -0.018 |
| | | (0.129) | (0.114) |
| Tuesday | | -0.009 | -0.009 |
| | | (0.127) | (0.112) |
| Wednesday | | 0.038 | 0.050 |
| | | (0.126) | (0.112) |
| Thursday | | 0.091 | 0.123 |
| | | (0.126) | (0.111) |
| Average max wave height | | | 0.091 |
| last 2 days wave height | | | (0.022)** |
| Average max wave height | | | 0.047 |
| 3 & 4 days lagged | | | (0.021)* |
| Constant | -0.051 | -0.073 | -0.920 |
| | (0.080) | (0.115) | (0.190)** |
| Observations | 97 | 97 | 97 |
| R-squared | 0.08 | 0.09 | 0.31 |

Standard errors in parentheses. *significant at 5%; ** significant at 1%.

# Multiple Choice Q 1

9. Suppose that you run a multiple regression analysis and you find that the p-value in the ANOVA section of the output (associated with the F statistic) is less than the level of significance $\alpha$. However, none of the p-values associated with the individual coefficients is less than $\alpha$. Which of the following is the most likely cause of such a result?
    a. The error terms have unequal variance
    b. The model is not linear
    c. Multicollinearity
    d. The error terms are not independent
    e. The overall relationship is not significant

# Multiple Choice Q 2

20. There is evidence that stock market prices are more likely to fall on cloudy days than on sunny days.

    A. The evidence proves that investors are irrational.

    B. If conventional wisdom is that asset price fluctuations should not be affected by cloudiness, there is evidence of an anomaly in asset prices.

    C. This evidence demonstrates that asset markets are weak-form ***in***efficient.

    D. This evidence demonstrates that asset markets are semi-strong form ***in***efficient but may be weak-form efficient.

# Multiple Choice Q 3

1. Which of the following time series have unit roots? (The 5% critical value for the DF/ADF test in each case is −2.89; the 1% critical value is −3.51).

$$
\begin{array}{rcccccccc}
\Delta X_t & = & 1.80 & - & 0.60\,X_{t-1}; & & & & \\
 & & (0.30) & & (0.20) & & & & \\
\Delta Y_t & = & 0.51 & + & 0.90\,Y_{t-1} & + & 0.20\,\Delta Y_{t-1}; & & \\
 & & (3.55) & & (0.30) & & (0.05) & & \\
\Delta Z_t & = & 0.53 & - & 0.25\,Z_{t-1} & + & 0.14\,\Delta Z_{t-1} & - & 0.10\,\Delta Z_{t-2}. \\
 & & (0.43) & & (0.50) & & (0.04) & & (0.04)
\end{array}
$$

Figures in parentheses are standard errors.

# Multiple Choice Q 4

18. In the Capital Asset Pricing Model (CAPM) the Security Market Line (SML) represents:

    A. The set of all mean-variance efficient portfolios in capital market equilibrium.

    B. The linear relationship between the risk premium on asset $j$, $(\mu_j - r_0)$, and the risk premium on the market portfolio, $(\mu_M - r_0)$.

    C. The linear relationship between each asset's expected rate of return, $\mu_j$, and its *beta-coefficient*, $\beta_j$.

    D. The linear relationship between each asset's *beta-coefficient*, $\beta_j$ and its *Sharpe ratio*, $s_j$.

Briefly describe the consequences of ignoring the following phenomena when using least squares as an estimation technique for a multiple regression model:

a) Heteroscedasticity in the error term

b) Serial Correlation in the error term

c) Near Multicollinearity among the explanatory variables

---

Consider the consumption function $C_t = \alpha + \beta Y_t + u_t$ where $C_t$ is aggregate consumption expenditures in the United States and $Y_t$ is disposable income, both in constant dollars. Suppose you have annual data from 1930 to 1997. Because the World War II years were different, you wish to exclude them from any model estimation procedure. Furthermore, suppose you believe the postwar consumption function is likely to be different from the prewar function. Describe how you would test for this using dummy variable(s).

# Answer Q 3

## Unit Roots and Stochastic Trends: Sketch Solutions

1. Recall that in the DF/ADF tests, the statistic of interest is the usual t-statistic for testing the null hypothesis of whether the coefficient on the lagged level of the variable (e.g. $X_{t-1}$) is zero. The alternative hypothesis is that the coefficient is negative, resulting in a one-sided test, the critical values for which are negative. Taking each variable in turn:
   $X$: $\hat{\tau}_\mu = -0.6/0.2 = -3$, and $-3.51 < -3 < -2.89$ so the statistic is significant at the 5% level (suggesting we reject a unit root) but not at the 1% level (suggesting we do not reject a unit root). The evidence against a unit root is not strong.
   $Y$: $\hat{\tau}_\mu = 0.9/0.3 = +3$ which is not significant i.e. $-2.89 < 3$ so we do not reject a unit root.
   $Z$: $\hat{\tau}_\mu = -0.25/0.50 = -0.5$, and $-2.89 < -0.5$, so the statistic is not significant and we do not reject a unit root.