

Department of Economics
University of Maryland

John C. Chao
Spring 2013

Economics 422
Final Examination

This exam contains 5 questions. You should try to do all 5 questions. Please provide explanations for all your answers. Right answers with no explanation will receive no credit. On the other hand, wrong answers with thoughtful explanations may receive some credit. Partial credit will be given. **GOOD LUCK!**

QUESTION 1 (10 POINTS)

Suppose that you wish to conduct an empirical analysis of the relationship between age and earnings, and you believe that the relationship is nonlinear in variables. Hence, you estimate the following third-order polynomial regression, using age-earnings data for 1744 workers and controlling for the effect of gender by using a binary variable that takes on the value of one for females and is zero otherwise:

$$\begin{aligned}\widehat{Earn} &= \underset{(120.13)}{-683.21} + \underset{(9.27)}{65.83} \times Age - \underset{(0.22)}{1.05} \times Age^2 + \underset{(0.002)}{0.005} \times Age^3 - \underset{(12.45)}{163.23} \times Female, & (1) \\ R^2 &= 0.225, \quad SER = 259.73,\end{aligned}$$

where $Earn$ is weekly earnings in dollars and where the numbers in parentheses are standard errors.

- (a) Based on this estimation result, calculate the effect on earnings of a change in age by one year from 25 to 26, holding constant the gender variable.

Suppose you decide to also estimate a fourth-order polynomial regression, and you obtain the following results

$$\begin{aligned}\widehat{Earn} &= \underset{(283.11)}{-795.90} + \underset{(29.29)}{82.93} \times Age - \underset{(1.06)}{1.69} \times Age^2 + \underset{(0.016)}{0.015} \times Age^3 - \underset{(0.0009)}{0.0005} \times Age^4 \\ &\quad - \underset{(12.45)}{163.19} \times Female, & (2) \\ R^2 &= 0.225, \quad SER = 259.78\end{aligned}$$

- (b) Based on the results obtained above, which specification would you prefer, (1) or (2)? Explain and provide argument in support your preference.

QUESTION 2 (24 POINTS)

Consider the Cobb-Douglas production function

$$Q_t = CL_t^\alpha K_t^\beta e^{u_t}, \quad t = 1, \dots, T \quad (3)$$

where Q_t denotes output at time period t (i.e., the real value of goods produced in time t), L_t is labor input at time t , and K_t is capital input at time t . C , α , and β are unknown parameters of the model, while u_t is a random term.

- (a) Suggest a transformation of the variables that will allow you to represent the model (3) as a linear regression model.
- (b) How would you interpret α and β here? Explain carefully.
- (c) State additional assumptions for OLS estimation of α and β to be unbiased and consistent.

QUESTION 3 (20 POINTS)

Consider the simple instrumental variable regression

$$\begin{aligned} Y_i &= \beta X_i + u_i, \\ X_i &= \pi Z_i + v_i, \quad i = 1, \dots, n; \end{aligned}$$

where (Y_i, X_i, Z_i) are *i.i.d.*; $Cov(u_i, v_i) \neq 0$; $\pi \neq 0$; and, with probability one, $E[u_i|Z_i] = 0$ and $E[v_i|Z_i] = 0$. Let

$$\hat{u}_i = Y_i - \hat{\beta}^{TSLS} X_i, \text{ for } i = 1, \dots, n,$$

where $\hat{\beta}^{TSLS}$ denotes the two-stage least squares (TSLS) estimator, so that \hat{u}_i is the TSLS residual. Now, consider the regression

$$\hat{u}_i = \gamma Z_i + \eta_i, \quad i = 1, \dots, n. \quad (4)$$

Show that

$$\hat{\gamma}^{OLS} = 0,$$

where $\hat{\gamma}^{OLS}$ denotes the ordinary least squares (OLS) estimator of γ in regression (4) above.

QUESTION 4 (30 POINTS)

Consider the simultaneous equations model

$$\begin{aligned} Q_i &= \beta_0 + \beta_1 P_i + \beta_2 Z_i + u_i, \\ P_i &= \pi_0 + \pi_1 Z_i + v_i, \end{aligned}$$

Suppose that (Q_i, P_i, Z_i) are *i.i.d.* and $Cov(u_i, v_i) \neq 0$. Suppose also that, with probability one, $E[u_i|Z_i] = 0$ and $E[v_i|Z_i] = 0$. Here, β_0 and π_0 are intercept parameters, and β_1 , β_2 , and π_1 are coefficient parameters.

- (a) Suppose that $\beta_2 \neq 0$, $\pi_0 \neq 0$, and $\pi_1 \neq 0$. Are the parameters β_0 , β_1 , and β_2 identified in this case? Can we estimate β_0 , β_1 , and β_2 consistently? Explain your answers carefully.
- (b) Alternatively, suppose that $\pi_0 = 0$, but $\beta_2 \neq 0$ and $\pi_1 \neq 0$. Now, which parameter(s), amongst β_0 , β_1 , and β_2 , can we identify? Which parameter(s) can we not identify? Which parameter(s), amongst β_0 , β_1 , and β_2 , can we estimate consistently. Explain your answers carefully.
- (c) Instead, suppose that $\pi_1 = 0$, but $\beta_2 \neq 0$ and $\pi_0 \neq 0$. Now, which parameter(s), amongst β_0 , β_1 , and β_2 , can we identify? Which parameter(s) can we not identify? Which parameter(s), amongst β_0 , β_1 , and β_2 , can we estimate consistently. Explain your answers carefully.

Finally, suppose the simultaneous equations model is of the form

$$\begin{aligned} Q_i &= \beta_0 + \beta_1 P_i + \beta_2 Z_i + u_i, \\ P_i &= \pi_0 + \pi_1 Z_i + \pi_2 W_i + v_i, \end{aligned}$$

where (Q_i, P_i, Z_i, W_i) are *i.i.d.*; $Cov(u_i, v_i) \neq 0$; and, with probability one, $E[u_i|Z_i, W_i] = 0$ and $E[v_i|Z_i, W_i] = 0$.

- (d) Suppose that $\beta_2 \neq 0$, $\pi_0 \neq 0$, $\pi_1 \neq 0$ and $\pi_2 \neq 0$. Are the parameters β_0 , β_1 , and β_2 identified in this case? Can we estimate β_0 , β_1 , and β_2 consistently? Explain your answers carefully.

QUESTION 5 (16 POINTS)

Suppose that you are the head of a consulting group for a particular fast food hamburger chain. Your group tries to determine sales revenue in different cities by estimating the following regression

$$SALES_i = \beta_0 + \beta_1 PRICE_i + \beta_2 ADEXP_i + \beta_3 (ADEXP_i)^2 + u_i \quad (5)$$

where $SALES_i$ represents the average monthly sales revenue from selling hamburgers in city i , $PRICE_i$ denotes the price charged for hamburger in city i , and $ADEXP_i$ is the average monthly expenditure on advertising in city i . Both $SALES_i$ and $ADEXP_i$ are measured in terms of thousands of dollars. Suppose that a brilliant member of your group has determined that the optimal amount of monthly advertising expenditure is given by the formula

$$\beta_2 + 2(\beta_3 \times ADEXP_{\text{opt}}) = 1,$$

The fast food chain currently spends \$3000 per month on advertising, and the management team for the chain would like you to determine if this amount of advertising expenditure is optimal, i.e., if it is indeed the case that $ADEXP_{\text{opt}} = 3$. Hence, you decide to test the null hypothesis

$$H_0 : \beta_2 + 2(\beta_3 \times 3) = \beta_2 + 6\beta_3 = 1 \quad \text{versus} \quad H_1 : \beta_2 + 6\beta_3 \neq 1$$

Show how you would transform or rearrange regression (5), so that you can test this hypothesis as a restriction on a single coefficient of the rearranged regression. Describe how you would go about testing this null hypothesis, by giving the test statistic to be used, the (exact or approximate) distribution of the test statistic under the null hypothesis, the critical value at the 5% significance level, and the decision rule for rejecting H_0 . For the purpose of answering this question, you can assume that regression (5) satisfies all the usual assumptions that we have made in class for a linear regression that is nonlinear in variables. (In addition, do not actually try to implement the test since of course you do not have sufficient information for that.)