Ph.D. Econometrics II
Heinz School, Carnegie Mellon University
90-907, Fall 2001

Final

Instructions You may use any books, notes, calculators, and other aids you like. You may not converse, nor may you cooperate.

Please complete all questions.

Questions one and four are worth 30 points each. Questions two and three are worth 20 points each.

Please show all relevant work.

Please interpret your results in plain English.

1. Consider a conditional logit model of consumer demand for automobiles:

$$U_{ij} = -\beta_1 \text{price}_j + \beta_2 \text{weight}_j + \epsilon_{ij} \tag{1}$$

Price is measured in $1000, and weight is measured in 1000 lbs. You have the following estimates:

$$\begin{pmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{pmatrix} = \begin{pmatrix} 0.4 \\ 0.8 \end{pmatrix} \tag{2}$$

$$V\begin{pmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{pmatrix} = \begin{bmatrix} 0.004 & 0.001 \\ 0.001 & 0.005 \end{bmatrix} \tag{3}$$

Your engineers report back to you that it will cost about $1 per pound to increase the car's size. Assume this estimator is distributed normal with a standard error of 0.2 and that it is independent of the demand estimators above.

You (the firm) are considering increasing the size of the car, meaning increasing the weight. Your strategy will be to maintain marketshare (so you will exactly offset any weight change by a price change).

(a) Please give an estimate and a 95% CI for the extra profit per car of the 1 lb weight increase.

(b) What is the highest cost which would yield a 90% confidence that the change would be profitable — assuming that the standard errors do not change?

2. Consider the following model (with scalar X and both X and Y zero-mean):

$$
\begin{aligned}
Y &= \beta X + \epsilon && (4)\\
\hat{X} &= X + \nu && (5)\\
\tilde{X} &= X + \gamma && (6)
\end{aligned}
$$

$$
\begin{aligned}
E\{\epsilon|X\} &= 0 & V\{\epsilon|X\} &= \sigma_\epsilon I && (7)\\
E\{\nu|X\} &= 0 & V\{\nu|X\} &= \sigma_\nu I && (8)\\
E\{\gamma|X\} &= 0 & V\{\gamma|X\} &= \sigma_\gamma I && (9)\\
& & & && (10)
\end{aligned}
$$

Furthermore, ϵ, γ, ν are independent of one another.

You observe Y, $\hat{X}$, $\tilde{X}$ but not X. This is much like a standard model of measurement error. If you just regress Y on $\hat{X}$ or on $\tilde{X}$, you will get a β-estimate biased towards zero.

(a) When is it better to regress Y on $\hat{X}$ and when is it better to regress Y on $\tilde{X}$?

(b) Devise some IV estimators of β, prove they are consistent and tell me which one is best.

3. Consider the classical linear regression model (with scalar X and zero-mean X and Y):

$$Y = \beta X + \epsilon \tag{11}$$

Obviously, the following equation is also true:

$$X = \frac{1}{\beta} Y + \frac{1}{\beta} \epsilon \tag{12}$$

As you saw on a previous year's exam, OLS on this equation does not produce a very useful estimator.

(a) What would happen if we ran IV on equation 12, using X as an instrument?

(b) How would that estimator be related to the OLS estimator on equation 11?

4. Consider the panel data model (again X a scalar and both X,Y zero-mean):

$$
\begin{aligned}
Y_{it} &= \beta X_{it} + \epsilon_{it} && (13) \\
\epsilon_{it} &= \alpha_i + \nu_{it} && (14) \\
X_{it} &= X_i + \gamma_{it} && (15)
\end{aligned}
$$

We are worried that X may be endogenous — that $Cov(X, \epsilon) \neq 0$. We will assume that $Cov(X_i, \nu_{it}) = Cov(\gamma_{it}, \alpha_i) = 0$ and $Cov(X_i, \gamma_{it}) = Cov(\alpha_i, \nu_{it}) = 0$, so that endogeneity may enter either through $Cov(X_i, \alpha_i)$ or through $Cov(\gamma_{it}, \nu_{it})$ or through both.

We are considering using either fixed effects (using unit but not time dummies) or OLS to estimate this equation — recall that, as far as consistency goes, random effects is essentially similar to OLS.

(a) Which estimation method, fixed effects or OLS, will result in the better estimator if only $Cov(X_i, \alpha_i) \neq 0$?

(b) Which estimation method, fixed effects or OLS, will result in the better estimator if only $Cov(\gamma_{it}, \nu_{it}) \neq 0$?

(c) Discuss what this means for using fixed effects to "get rid of" endogeneity in a regression.