

Econometrics (60 points)

Question 7: Short Answers (30 points)

Answer parts 1-6 with a brief explanation.

1. Suppose the model of interest is $Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + u_i$, where $E(u|X)=0$ and $E(u^2|X)=\sigma^2$ and X_1 and X_2 are uncorrelated in your sample. Will the bivariate regression of Y on X_i have the same coefficient estimate and standard error for $\hat{\beta}_1$ as the multivariate regression of Y on X_1 and X_2 ? [6 points]

Answer: The coefficient estimates will be the same but the standard error will be smaller in the multivariate regression. Let β_1^b be the coefficient on X_1 in the bivariate regression. Using the formula for omitted variables bias, we know that

$$\beta_1^b = \frac{Cov(Y, X_1)}{Var(X_1)} = \frac{Cov(\beta_1 X_1 + \beta_2 X_2 + u, X_1)}{Var(X_1)} = \beta_1 + \frac{\beta_2 Cov(X_2, X_1)}{Var(X_1)} + \frac{Cov(u, X_1)}{Var(X_1)}$$

By assumption, $Cov(X_2, X_1) = 0$ and $Cov(u, X_1) = 0$ and so $\beta_1^b = \beta_1$.
The value of β_1 in the multivariate regression can be written as

$$\beta_1^m = \frac{Cov(Y, \widetilde{X}_1)}{Var(\widetilde{X}_1)}$$

where $\widetilde{X}_1 = X_1 - \gamma_0 - \gamma_1 X_2 = X_1 - \overline{X}_1$ since $Cov(X_2, X_1) = 0$. Thus

$$\beta_1^m = \frac{Cov(Y, \widetilde{X}_1)}{Var(\widetilde{X}_1)} = \frac{Cov(\beta_1 X_1 + \beta_2 X_2 + u, X_1 - \overline{X}_1)}{Var(X_1 - \overline{X}_1)} = \beta_1.$$

The standard error of the vector β is given by $\sigma^2(X'X)^{-1}$ where σ^2 is the variance of the error term and X is the vector of independent variables. Including X_2 reduces the standard error on the estimate of β_1 because it reduces σ^2 but leaves the relevant term of $(X'X)^{-1}$ unchanged since X_1 and X_2 are uncorrelated.

Point Values:

- 2 points: The estimates will be the same
- 2 points: A reason
 - 1 point for a reasonable reason
 - 1 point if it includes any mathematical derivation
- 1 point: The standard error will decrease
- 1 point: Some reasonable reason

Parts 2 to 5 refer to the demand curve,

$$\ln(Q_t) = \beta_0 + \beta_1 \ln(P_t) + \beta_2 \ln(Y_t) + u_t, \quad (1)$$

where Q_t and P_t are the quantity (number) and price of haircuts obtained in Cambridge in year t and Y_t is mean income in Cambridge in year t .

2. Express the price elasticity of demand in terms of the coefficients in (1). [6 points]

Answer: The price elasticity of demand is β_1 , which is the derivative of $\ln(Q_t)$ with respect to $\ln(P_t)$.

Suppose you have annual data on Q_t , P_t , and Y_t in Cambridge for 30 years, and that you have some other annual data available too. You are interested in estimating the coefficients of equation (1). Assume price and quantity are simultaneously determined in a market equilibrium. Would the following variables plausibly be valid instruments for $\ln(P_t)$?

3. $\ln(Y_t)$. [6 points]

Answer: This is not valid. If the econometrician does not control for $\ln(Y_t)$, then the instrument is not exogenous, since clearly $\text{corr}(\beta_2 \ln(Y_t) + u_t, \ln(Y_t)) \neq 0$. If the econometrician does control for $\ln(Y_t)$, then the instrument is collinear with the controls.

Points: 1.5 points for the instrument not being valid
1.5 points for it not being exogenous
3 points for a reason

4. Commercial rental price (dollars/square foot/month) in Cambridge in that year. [6 points]

Answer: I could see an argument either way for this one. The instrument is plausibly relevant, since an increase in commercial rents could lead to haircutting salons increasing prices of haircuts even controlling for mean income. I can see an argument either way for exogeneity:

It is plausibly exogenous controlling for mean income, since controlling for mean income controls for general economic conditions which might impact both commercial rents and demand for haircuts.

It is not necessarily exogenous. Suppose that all the major investment banks decide to move from New York to Cambridge. Then this might cause an increase in commercial rents due to increased demand for space and also an increase in demand for haircuts as individuals shift from working at biotech startups (and having long hair) to working at investment banks.

Points: 1 point for valid or not
1 point for relevance
1.5 points for the reason why it is relevant
1 point for whether it is exogenous
1.5 points for the reason

5. The average length of hair of individuals appearing in *People* magazine in that year. [6 points]

Answer: This is not a valid instrument. It is not correlated with the price of haircuts and is correlated with unobservables that impact the demand for haircuts.

Point Values: 1 point for not valid
1.5 points for it not being relevant
1.5 points for it not being exogenous
2 points for reasoning

Question 8: Performance-linked pay (30 points)

Table IV from Lemieux, MacLeod, and Parent (*Quarterly Journal of Economics*, 2009; see the following page) shows results from a regression of log wages on a dummy for whether a job has pay linked to performance (e.g. salespeople paid on commission) and other variables. The data are panel data on workers. In addition to the reported coefficients, the regressions include industry, occupation, and year dummies; county unemployment; and marital status, race dummies, and union status. Standard errors are in parentheses.

The model also includes quadratic functions of experience (number of years in the workforce) and tenure (number of years at this specific job). The row labeled “Experience x performance-pay” is the effect of experience at 20 years interacted with performance pay. Similarly, the row labeled “Tenure x performance pay” is the effect of tenure (evaluated at ten years) interacted with performance pay.

1. Based on column (3), is the return to education higher at performance pay jobs or non-performance pay jobs? What is the difference and is it statistically significant? [6 points]

Answer: The returns to education are higher in performance pay jobs. The coefficient on education x performance pay is 0.0365 with a standard error of 0.007. The t statistic for the test that the coefficient is equal to zero is 5.214 which has a p value of 0.000.

Points: 1 point for higher

1 point for the coefficient

1 point for the standard error

1 points for the t statistic

1 point for the p value

1 point for stating it is statistically significant at 5% or 1%.

2. Again using column (3), what is the return to having a performance pay job for somebody with a college degree (16 years of education), 20 years of experience, and 10 years of tenure? [6 points]

Answer: The difference in log wages between a performance pay job and a non performance pay job for a person with the given characteristics is given by

$$\begin{aligned} & \beta_{pp} + 16 * \beta_{\{ed,pp\}} + 1 * \beta_{\{exp,pp\}} + 1 * \beta_{\{ten,pp\}} \\ & = -.4526 + 16 * 0.0365 + 1 * 0.1162 + 1 * -0.0666 = 0.617 \end{aligned}$$

The person will earn a staggering 61.7% more in log wages in a performance pay job than not.

Points: 1 point for β_{pp}

1 point for $16 * \beta_{\{ed,pp\}}$

1 point for $1 * \beta_{\{exp,pp\}}$

1 points for $1 * \beta_{\{ten,pp\}}$

1 point for getting the correct sum

1 point for interpreting the answer.

3. Regression (4) includes worker-level fixed effects. The coefficient on years of education falls from .0637 in (3) to .0167 in (4). Is this a large change in economic terms? Explain. [6 points]

Answer: This is a large effect. A coefficient of 0.06 indicates that, all other things equal, a one year increase in education increases expected log wage by 6%. A coefficient of 0.0167 indicates that an additional year of education increases expected wage by less than 2%, which is less than a third of the first effect. The wage returns to education appear much smaller in this second estimate. In addition, the coefficient is no longer statistically significant at 5%.

Points: 6 points for reasonable reasoning

4. Provide an explanation for the difference in the coefficients discussed in question 3 (.0637 vs. .0167). Be concrete. [6 points]

Answer: Including fixed effects controls for individual unobservables which would impact both educational attainment and log wages. For example, suppose that more able workers get more education and also are hired by more productive firms (and have higher wages as a consequence). Then regression 3 would suffer from omitted variables bias and this would lead to the coefficient on education being biased upward. Including individual fixed effects eliminates this OVB problem.

Points: 2 points for mentioning unobservables or omitted variables bias

2 points for giving an example of an omitted variable

2 points for explaining why this example variable would bias the coefficient upward

5. Consider three possible ways to compute standard errors for the regressions in Table IV: homoskedasticity-only; heteroskedasticity-robust; and clustered at the individual-job level. Which is the most appropriate method, and why? [6 points]

Answer: Standard errors should be clustered at the individual-job level. This is because even controlling for individual and industry characteristics, idiosyncratic portions of wages (such as firm-specific shocks that impact wages) may be correlated for an individual over time. Thus the error terms are not iid and the standard errors need to be corrected for this fact.

Points: 1 point for not picking homoscedastic

1 point for a reasonable explanation if they pick heteroscedastic

2 points for picking individual-job level

3 points for an explanation of why individual-job clustering is required

TABLE IV
SKILLS-RELATED WAGE DIFFERENTIALS AND PERFORMANCE-PAY (PP) JOBS

Estimation method	Sample					
	PP jobs OLS (1)	Non-PP jobs OLS (2)	All jobs			
			OLS (3)	FE (4)	OLS (5)	FE (6)
Performance-pay job dummy	—	—	−0.4526 (0.1019)	−0.2061 (0.0723)	−0.2406 (0.1251)	0.1414 (0.0998)
Years of education	0.0929 (0.0071)	0.0665 (0.0039)	0.0637 (0.0040)	0.0167 (0.0091)	0.0584 (0.0047)	0.0040 (0.0096)
Education × performance-pay job	—	—	0.0365 (0.0071)	0.0169 (0.0048)	0.0217 (0.0092)	−0.0079 (0.0071)
Education × 1990–1993	—	—	—	—	0.0161 (0.0085)	0.0222 (0.0056)
Education × performance-pay job × 1990–1993	—	—	—	—	0.0190 (0.0137)	0.0280 (0.0089)
Potential experience (effect at 20 years)	0.4259 (0.0535)	0.2882 (0.0288)	0.3010 (0.0294)	0.4545 (0.1258)	0.3002 (0.0294)	0.4231 (0.1256)
Experience × performance-pay job	—	—	0.1162 (0.0584)	0.0149 (0.0501)	0.1018 (0.0581)	−0.0278 (0.0509)
Tenure (effect at ten years)	0.1670 (0.0268)	0.2197 (0.0154)	0.2262 (0.0154)	0.1158 (0.0129)	0.2271 (0.0154)	0.1191 (0.0129)
Tenure × performance-pay job	—	—	−0.0666 (0.0301)	0.0278 (0.0237)	−0.0677 (0.0303)	0.0196 (0.0239)
Number of observations	9,680	16,466	26,146	26,146	26,146	26,146