# ECON 2223B: Homework 5

Due Saturday, April 11 at 11:59pm.

## Problem 1

*Hypothesis testing gives us a way to check whether data are consistent with an economic model. The idea is to take the economic model seriously: if it is correct, it places restrictions on what the regression coefficients should look like. We then test those restrictions using our estimates. Rejecting the restriction means the data are not consistent with the model, i.e., we'd reject the model. (Failing to reject does not prove the model is correct. Rather, it just means the data are consistent with it.) This problem asks you to carry out that logic for a supply-side model.*

A manufacturing firm has Cobb-Douglas production technology:

$$Q = A \cdot K^{1-\tilde{\alpha}} \cdot L^{\tilde{\alpha}}, \qquad \tilde{\alpha} \in (0,1),$$

where $Q$ is output, $L$ is labour hired, $K$ is capital, $A > 0$ is productivity, and $\tilde{\alpha}$ is the **labour elasticity of output**. The firm takes the output price $P$ and the market wage $w$ as given, and chooses how much labour $L$ to hire.

The first-order condition for optimal labour demand is $P \cdot \partial Q / \partial L = w$. For the Cobb-Douglas technology above, $\partial Q / \partial L = \tilde{\alpha} \cdot Q/L$, so the condition can be rearranged to the **labour demand equation**:

$$L^* = \tilde{\alpha} \cdot \frac{P \cdot Q}{w}.$$

The firm's optimal labour input is proportional to its revenue $PQ$ relative to the wage $w$, with the proportionality constant equal to $\tilde{\alpha}$.

The optimal labour input $L_i^*$ is not directly observed. However, the analyst observes:

$$L_i = L_i^* + u_i,$$

where $u_i$ is measurement error in labour. The analyst observes a random sample of $n$ firms $\{w_i, P_i, Q_i, L_i\}_{i=1}^n$.

i. Suppose the analyst observes the random sample above, where $n$ is large. Write down a statistical model suitable for OLS estimation of $\tilde{\alpha}$. What is the dependent variable and what is the regressor? Include an intercept and slope (suggested notation: $\beta_0$ for the intercept, $\beta_1$ for the slope), and relate $\beta_0$ and $\beta_1$ to $\tilde{\alpha}$.

   *Hints:*

   - *Note that $\tilde{\alpha}$ is purposely not subscripted. Part of the question is relating $\beta_0$ and $\beta_1$ to $\tilde{\alpha}$.*
   - *Consider generating a new variable from the data.*

ii. State clearly what assumptions are needed for the OLS estimator of $\beta_1$ to be consistent for $\tilde{\alpha}$.

iii. Under the Cobb-Douglas production model, what should the intercept $\beta_0$ and slope $\beta_1$ of your statistical model from part (i) be? Express in terms of $\tilde{\alpha}$, where relevant.

iv. Assume conditions for consistent estimates are met. The analyst estimates the statistical model from part (i) and obtains $\hat{\beta}_1 = 1.30$ with s.e.$(\hat{\beta}_1) = 0.14$. Using your answer from part (iii), test whether this finding is consistent with Cobb-Douglas production at the 5% significance level. State $H_0$ and $H_1$, compute the test statistic, and state your conclusion.

*Hint: think about what range of values the Cobb-Douglas specification allows for $\beta_1$. Does $\hat{\beta}_1 = 1.30$ fall within this range? What is the appropriate null to test, and why?*

## Problem 2

*Estimating a demand curve from market data is not as straightforward as it might seem. The prices we observe are **equilibrium** prices, i.e., they are set by the simultaneous interaction of supply and demand. This means observed prices are correlated with demand shocks, and OLS on the demand equation picks up a mixture of supply and demand effects. Instrumental variables provide a way out: if we can find a variable that shifts supply without shifting demand, it generates variation in price that traces out the demand curve. This problem asks you to work through that logic in the context of the rental housing market.*

A housing economist wants to estimate the demand curve for rental apartments. The economic models for supply and demand are

$$Q^D = \tilde{\delta}_0 + \tilde{\delta}_1 P + \varepsilon^D, \qquad Q^S = \tilde{\gamma}_0 + \tilde{\gamma}_1 P + \varepsilon^S,$$

where $P$ is monthly rent (dollars) and $Q$ is units rented. In equilibrium, $Q^D = Q^S$ and the market clears at the observed rent. The analyst is interested in the demand slope parameter $\tilde{\delta}_1 < 0$.

The analyst has data from a random sample of cities and specifies the statistical model for the demand side:

$$Q_i = \beta_0 + \beta_1 P_i + u_i, \qquad \beta_1 = \tilde{\delta}_1.$$

i. Write down the formula for the OLS estimator $\hat{\beta}_1$ of $\tilde{\delta}_1$. Briefly justify where this formula comes from.

ii. Would OLS consistently estimate $\tilde{\delta}_1$? Explain why or why not, referencing your answer to part (i).

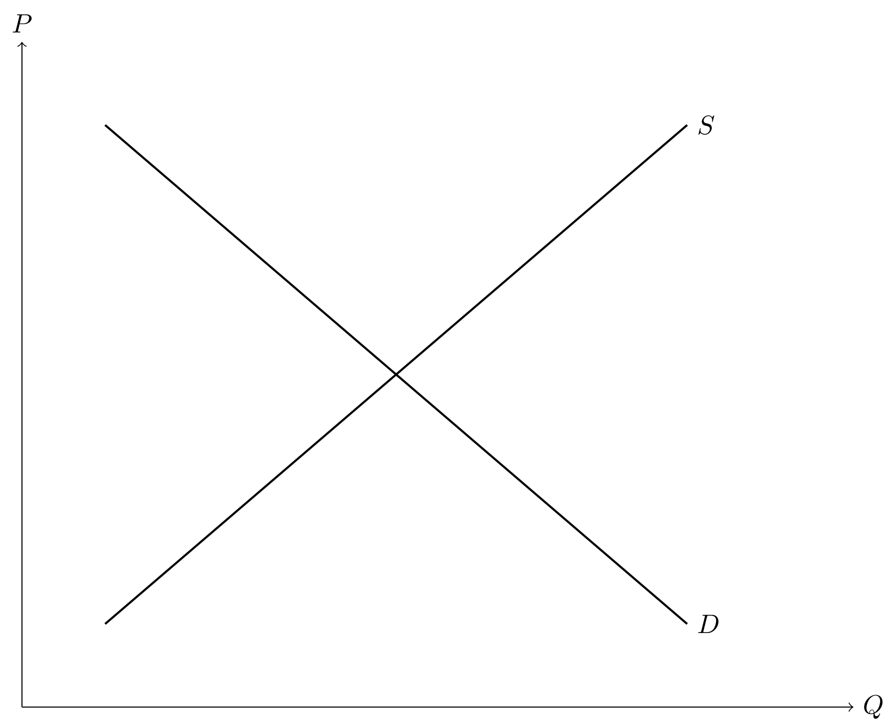
iii. The analyst considers five potential instruments for $P_i$. For each, discuss whether it would serve as a valid instrument to obtain consistent estimates of $\tilde{\delta}_1$.

*Hint: What are the two conditions an instrumental variable must satisfy to deliver consistent estimates?*

(a) Average household income in the city.
(b) National prices for lumber and concrete (construction input costs).
(c) A new city-wide zoning bylaw permitting higher-density residential construction.
(d) The opening of a large tech-company headquarters in the city.
(e) Mortgage interest rates.

iv. Of the conditions required for a valid instrument, which are testable using data, and which must be maintained as untestable assumptions? Explain.

v. Suppose the analyst uses a valid instrument from part (iii). Explain the steps of the IV procedure: what would the analyst regress on what in each stage? Then, using the supply-and-demand diagram below, illustrate how a valid instrument identifies $\tilde{\delta}_1$. Note: the diagram below is drawn in standard economist $(Q, P)$ space ($Q$ on the horizontal axis), so the geometric slope of the demand curve as drawn is $\Delta P/\Delta Q = 1/\tilde{\delta}_1$, not $\tilde{\delta}_1$ itself. Consider two values of the instrument, $Z_0 < Z_1$. Label the resulting equilibria $(Q_0, P_0)$ and $(Q_1, P_1)$. Show which curve shifts, in which direction, and how the two equilibrium points identify $\tilde{\delta}_1$.




## Problem 3

A researcher wants to estimate the causal effect of exercise on health. She has panel survey data on individuals over two time periods ($t = 1, 2$) and specifies the model

$$\text{health}_{it} = \beta_0 + \beta_1 \,\text{exercise}_{it} + a_i + u_{it},$$

where $\text{health}_{it}$ is a standardised health index for individual $i$ in period $t$, $\text{exercise}_{it}$ is weekly hours of exercise, and $a_i$ is an unobserved time-invariant individual characteristic.

i. What might $a_i$ represent in this context? Give a specific example. Explain why pooled OLS (regressing $\text{health}_{it}$ on $\text{exercise}_{it}$, ignoring $a_i$) might not deliver a consistent estimate of $\beta_1$.

ii. Using the omitted variable bias formula, express the probability limit of the pooled OLS estimator of $\beta_1$ in terms of $\beta_1$ and the relationships between $\text{exercise}_{it}$ and each of $a_i$ and $u_{it}$. In which direction is the OLS estimate of $\beta_1$ likely biased? Explain your reasoning.

iii. Describe a method to estimate $\beta_1$ that eliminates $a_i$. Write the equation that justifies your method, starting from the equations for $t = 1$ and $t = 2$. What would you regress on what? State the key assumption needed for this estimator to consistently estimate $\beta_1$.

iv. Explain why the assumption from part (iii) might fail in this setting, even after $a_i$ has been removed. Give a specific example of the mechanism, and state the likely direction of bias.

Between the two survey periods, a new fitness facility opened near the homes of some individuals. Define $Z_i = 1$ if a new fitness facility opened within 2 km of individual $i$'s home between $t = 1$ and $t = 2$, and $Z_i = 0$ otherwise.

v. For $Z_i$ to be a valid instrument for $\Delta\text{exercise}_i$ in the first-differenced model, two conditions must hold. State each condition and discuss whether it is plausible in this setting. Which condition is testable? Then write an expression for what $\hat{\beta}_1^{IV}$ converges to. Is the IV estimator consistent for $\beta_1$? Explain.

vi. A colleague argues: "Fitness facilities tend to open in neighbourhoods that are becoming wealthier and more health-conscious — places where residents' health is improving for other reasons." Which IV condition does this threaten? How would it affect the probability limit of $\hat{\beta}_1^{IV}$?