

WORKSHOP

This workshop does not involve any computer work: you get a break till tomorrow morning! Tomorrow you will do IV and 2SLS estimation in the context of simultaneous equations models. Today we will examine a variety of situations where OLS is biased, exploring the direction of bias and the choice of instruments to perform unbiased IV/2SLS.

I.

Recall Monday, when you estimated the effect of output level and capacity on the average cost of an electricity generating facility:

$$AC = \alpha + \beta Q + \theta K + \varepsilon$$

An energy utility generally operates several plants and dispatches the plants in a cost-conscious manner. Plants with low production costs are generally used more than high-cost plants, and high-cost plants are used only when total demand is sufficiently great to require them.

Under this form of dispatching, there are two directions of causation between Q and AC .

- (1) Q causes AC : producing more output in a plant lowers the average cost (holding capacity constant). This is the effect that we are trying to estimate in β .
 - (2) AC causes Q : plants with lower average costs (for reasons other than their output--such as because they are newer and more efficient) are used more and hence have higher output.
1. What direction of bias does this form of dispatching produce in the OLS estimate of β ? That is, can we expect the OLS $\hat{\beta}$ to be larger (in magnitude) than the true β , or smaller?

2. What instruments would be appropriate to use in IV estimation? An appropriate instrument is a variable that relates to the quantity produced but does not have an independent effect on average costs (i.e., is not an element of ε). For each of the following variables, decide whether the variable could serve as an instrument.

- Weather variables. For example: cooling degree days in 1992 in the service territory of the utility that owns and operates the facility. Weather affects the total demand that the utility must meet and hence affects the amount each facility is used by the utility.
- Reserve margin of the utility that owns and operates the facility. The utility's reserve margin is $1 - (TQ/TK)$ where TQ is peak megawatts of the utility, including all its facilities, and TK is the total capacity. A low reserve margin means the utility must generally use each of its facilities more.
- Measures of economic activity in the service territory of the utility that owns the facility, such as average income, employment rates, etc. These variables, like the weather variables, affect the total demand that the utility must meet.
- Average price of electricity sold by the utility that owns the facility. Like weather and economic activity, price affects the total demand that the utility must meet. Assume, as is usual, that the utility is under rate-of-return regulation, such that its price depends on its costs.
- Age of the facility (that is, how long ago it was built).
- Average wages in the county in which the facility is located.

Can you think of any other good instruments?

3. Suppose you have collected data on variables that you think serve as good instruments. Label these variables IONE and ITWO. Write the TSP commands that will estimate your model by 2SLS, with the two steps performed explicitly. Then write the TSP commands that will estimate your model by IV. (If you ran these two programs, you would get the same results.)

II.

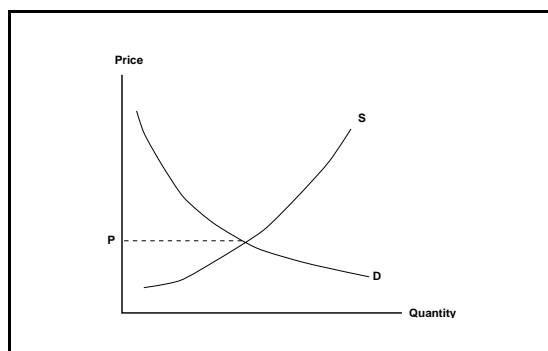
Suppose you want to estimate the price elasticity of demand for a good that is sold in a competitive market--like salmon. You observe the total quantity of salmon sold monthly in the Bay Area and the price of salmon each month, for a period of ten years.

You want to estimate the demand curve:

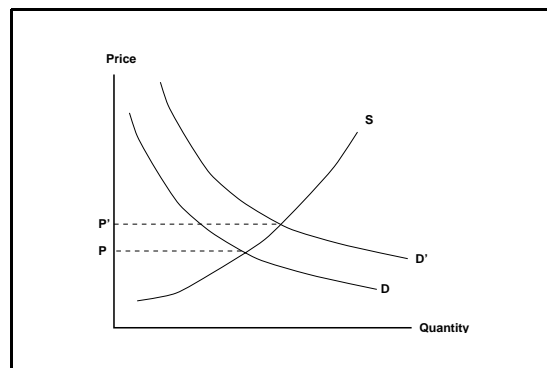
$$QD = \alpha - \beta P + \varepsilon .$$

You expect demand to be price-responsive: when price rises, the quantity demanded drops. The true β captures the magnitude of this response; you want to estimate β .

You realize, however, that salmon is sold in a competitive market, in which price is determined by the interaction of supply and demand:



Therefore, when the quantity demanded rises for some non-price reason (e.g., new evidence that beef can be unhealthy induces people to switch to fish), the price of salmon rises:

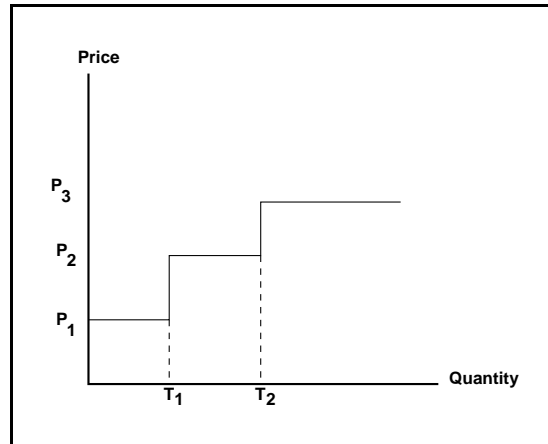


There are two directions of causation between QD and P. (1) P causes QD: when price rises, the quantity demanded falls. This is the effect you want to estimate in β . (2) QD causes P: when the quantity demanded rises for a non-price reason, then price rises (due to the interaction of supply and demand).

1. Which direction of bias does the interaction of supply and demand cause in the OLS estimate of β ?
2. What variables would serve as good instruments in an unbiased IV/2SLS estimation?

III.

You want to estimate the price-elasticity of demand for electricity using customer-level data. You realize, however, that electricity is priced in blocks:



A customer who consumes in tier 1 (that is, less than T_1 kWh) faces a price of P_1 . A customer who consumes in the second tier pays P_1 for its first T_1 units and faces a price of P_2 for each unit above T_1 . Similarly for a customer consuming in tier 3.

Because of these block rates, a distinction is made between the average and marginal price faced by the customer. The average price is the customer's bill divided by its total consumption. The marginal price is the price in the tier in which the customer consumes (i.e., the price of an extra unit). For a customer in the first tier, average and marginal price are the same: P_1 . For a customer in the second tier, marginal price is P_2 and average price is an average of P_1 and P_2 . (To be exact, average price is $P_1(T_1/Q) + P_2((Q - T_1)/Q)$ where Q is total consumption.) A customer in the third tier has P_3 as its marginal price, and its average price is an average of P_1 , P_2 , and P_3 .

You obtain the energy bills for a sample of households in a given month. From the bill, you observe the household's total electricity consumption for the month (Q), the marginal price that the household faces for electricity (MP), and the total charge for electricity (R). You calculate the household's average price of electricity as $AP = R/Q$. You also collect data on each household's income, size, dwelling type, the weather in its area, and other factors that affect households' electricity consumption. Label these non-price variables as X_1, X_2, \dots, X_N .

You want to estimate a model of households' electricity consumption, including the response to price. That is, you want to estimate:

$$Q = \alpha + \beta(\text{Price}) + \lambda_1 X_1 + \dots + \lambda_N X_N + \varepsilon.$$

You have a choice of using MP or AP for price. Economic theory suggests that consumers respond to marginal price, but anecdotal concepts indicate that consumers respond to average price. Whichever price variable you choose, the same issues arise because of the block-rate prices for electricity.

1. Explain why the OLS estimate of β (using either MP or AP) is biased. Determine the direction of bias.
2. A very convenient method for correcting this bias has been developed for this situation. The procedure is:
 - Regress Q against all the explanatory variables **except** price. That is, estimate

$$Q = \alpha + \lambda_1 X_1 + \lambda_2 X_2 + \dots + \lambda_N X_N + \varepsilon^*$$

- Calculate a predicted Q from this regression, labeled \hat{Q} .

- Determine the tier that \hat{Q} falls in. The price in this tier is called the "instrumented marginal price," labeled IMP. Note that if \hat{Q} falls in a different tier than Q then IMP will be different from MP--that is, the instrumented marginal price will be different than the observed marginal price.

Similarly, calculate the bill that the household would receive if it consumed \hat{Q} . Dividing this "bill" by \hat{Q} gives the "instrumented average price," labeled IAP.

- Estimate the original regression with IMP or IAP as the price variable (depending on whether you want to capture response to marginal or average price).

Why does this procedure provide an unbiased estimate of the price response?