useful for estimating underlying properties.

$$q_i^S = \alpha_2 p_i + \varepsilon_i^S$$
$$q_i^D = \beta_2 p_i + \beta_3 y_i + \varepsilon_i^D$$
$$q_i^D = q_i^S$$

structural form equations

y_i is exogenous $\Leftrightarrow E[y_i \varepsilon_i^S] = 0$, $E[y_i \varepsilon_i^D] = 0$

reduced form equations

$$q_i = \frac{\alpha_2 \beta_3}{\alpha_2 - \beta_2} y_i + \frac{\alpha_2 \varepsilon_i^D - \beta_3 \varepsilon_i^S}{\alpha_2 - \beta_2} = \pi^q y_i + \varepsilon_i^q$$
$$p_i = \frac{\beta_3}{\alpha_2 - \beta_2} y_i + \frac{\varepsilon_i^D - \varepsilon_i^S}{\alpha_2 - \beta_2} = \pi^p y_i + \varepsilon_i^p$$

solely useful for forecasting
- you don't need to know economics for this

What if we wanted to estimate the effects on prices and quantities of a change in sales tax? Use the structural form derived from economic theory to pin down price elasticities.
- How do we do this if we can't separate q_i^S and q_i^D?

Question of identification: Can we use π^q and π^p estimates to pin down α_2, β_2, and β_3?
- In general, no. In this case,

$\frac{\pi^q}{\pi^p} = \alpha_2$ by dumb luck. This method is not systematic.

Is there a systematic approach?
Can we always identify α_2, β_2, and β_3?

What is the nature of this $\frac{\pi^q}{\pi^p} = \alpha_2$ estimator?

$$\hat{\alpha}_2 = \frac{\hat{\pi}^q}{\hat{\pi}^p} = \frac{\left(\sum_{i=1}^n y_i q_i\right)/\left(\sum_{i=1}^n y_i^2\right)}{\left(\sum_{i=1}^n y_i p_i\right)/\left(\sum_{i=1}^n y_i^2\right)} = \frac{\sum_{i=1}^n y_i q_i}{\sum_{i=1}^n y_i p_i}$$

This is an IV estimator.

ILS ← IV

OLS estimators

∟ This is called the indirect least squares estimator

∟ Indirect least squares is numerically equivalent to IV estimator

Old-timer motivation ① $\hat{p}_i = \hat{\pi}^p y_i$ ← fitted values (regress p on instrument)

② Regress q_i on $\hat{p}_i$ $\quad q_i = \alpha \hat{p}_i + \varepsilon_i$

$$\Rightarrow \alpha = \frac{\sum_{i=1}^n \hat{p}_i q_i}{\sum_{i=1}^n \hat{p}_i^2} = \frac{\sum_{i=1}^n (\hat{\pi}^p y_i) q_i}{\sum_{i=1}^n (\hat{\pi}^p y_i)^2} = \frac{1}{\hat{\pi}^p} \frac{\sum_{i=1}^n y_i q_i}{\sum_{i=1}^n y_i^2}$$

regression of q on y

$$= \frac{1}{\hat{\pi}^p} \hat{\pi}^q$$

This is called the 2 stage least squares estimator. (written 2SLS)

ILS = IV = 2SLS

Can we use y_i as an estimate to get β_2?

For the first equation

$E[y_i \varepsilon_i^s] = 0$

$\Leftrightarrow E[y_i q_i] = \alpha_2 E[y_i p_i]$

$$\Leftrightarrow \alpha_2 = \frac{E[y_i q_i]}{E[y_i p_i]} \approx \frac{\frac{1}{n}\sum_{i=1}^n y_i q_i}{\frac{1}{n}\sum_{i=1}^n y_i p_i}$$ ← IV estimator

Can we apply this same logic to equation 2?

$E[y_i \varepsilon_i^D] = 0$

$\Leftrightarrow E[y_i q_i] = \beta_2 E[y_i p_i] + \beta_3 E[y_i^2]$

This is one equation in two unknowns. Identification fails. any information about β_3 will be enough to allow us to pin down β_2.

Suppose

$q_i^S = \alpha_2 p_i + \alpha_3 y_i + \varepsilon_i^S$

Here, we cannot identify α_2 using ILS.

Thus, we need $\alpha_3 = 0$. "zero restriction" "exclusion restriction."

Simultaneity need not arise from equilibrium conditions.

Causality of advertising and sales. (Nerlove, Waugh)

q_t = quantity

p_t = price

y_t = income of consumers

a_t = advertising.

Wanted to find effect of a_t on $p_t q_t$

$$q_t = k\, p_t^{\eta}\, y_t^{\beta}\, a_t^{\overset{\text{current effect}}{\gamma}}\, A_t^{\overset{\text{lagged effect}}{\delta}}$$

$$A_t = \frac{1}{10}(a_{t-1} + \dots + a_{t-10})$$

Taking logs:

$$\log q_t = \log k + \eta \log p_t + \beta \log y_t + \gamma \log a_t + \delta \log A_t$$

$$\Rightarrow \log p_t = c + \frac{1}{\eta}\log q_t - \frac{\beta}{\eta}\log y_t - \frac{\gamma}{\eta}\log a_t - \frac{\delta}{\eta}\log A_t$$

where $c \equiv -\frac{1}{\eta}\log k$

define revenue: $v_t = p_t q_t$

$$\Rightarrow \log v_t = \log p_t + \log q_t$$

$$= c + \left(1 + \frac{1}{\eta}\right)\log q_t - \frac{\beta}{\eta}\log y_t - \frac{\gamma}{\eta}\log a_t - \frac{\delta}{\eta}\log A_t$$

OLS gives us $\widehat{\left(1 + \frac{1}{\eta}\right)} = -0.390$

$$\widehat{\left(-\frac{\beta}{\eta}\right)} = 0.924$$

$$\left(-\frac{\gamma}{\eta}\right) = 0.233$$

$$\left(-\frac{\delta}{\eta}\right) = 0.103$$

$$\Rightarrow \hat{\eta} = -0.72$$

$$\hat{\beta} = 0.67$$

$$\hat{\gamma} = 0.17$$

$$\hat{\delta} = 0.07$$

Is this right? There is a measurement issue (cheap criticism)

Decisionmaking of marketing people.

- advertising affects sales
- sales affect advertising

$\}$ endogeneity

We assumed every variable was exogenous, which is not necessarily true.

Perhaps the correct model is given by:

$S_t \equiv$ sales

$M_t \equiv$ advertising budget

$P_{s,t} \equiv$ price of output

$P_{m,t} \equiv$ price of marketing

$$\left.\begin{aligned} S_t &= a + b \cdot M_t + c\, P_{s,t} + u_t \\ M_t &= d + e \cdot S_t + f\, P_{m,t} + v_t \end{aligned}\right\}$$

Note: simultaneity is not driven by equilibrium conditions as before

Running a regression of one endogenous variable on another endogenous variable is not okay

Identification Problem

(ignoring the i or t subscripts)

G equations

G endogenous variables

K exogenous variables

error terms

$$\begin{cases} y_1 + \beta_{12} y_2 + \dots + \beta_{1G} y_G + \gamma_{11} x_1 + \dots + \gamma_{1K} x_K + u_1 = 0 \\ \beta_{21} y_1 + y_2 + \dots + \beta_{2G} y_G + \gamma_{21} x_1 + \dots + \gamma_{2K} x_K + u_2 = 0 \\ \vdots \\ \beta_{G1} y_1 + \beta_{G2} y_2 + \dots + y_G + \gamma_{G1} x_1 + \dots + \gamma_{GK} x_K + u_G = 0 \end{cases}$$

Can we identify all these β's and γ's?

We have $G^2 - G$ β's

GK γ's