Instrumental Variables

Monday, November 8, 2021 3:23 PM

Commodity markets with supply + demand

$$\begin{cases} q_{st} = \alpha + \beta p_t + u_t \\ q_{dt} = c - d p_t + v_t \\ q_{st} = q_{dt} \end{cases}$$

$\alpha + \beta p_t + u_t = c - d P_t + v_t$

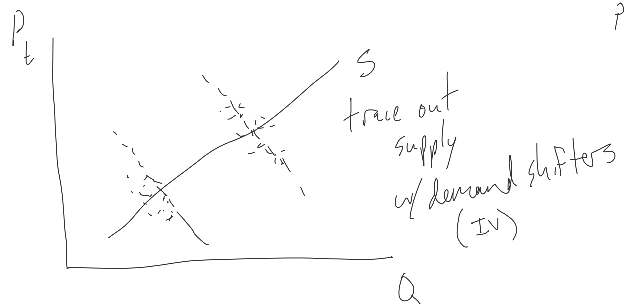
solve for P_t:

$$P_t = \frac{c - \alpha}{\beta + d} + \frac{v_t - u_t}{\beta + d}$$

$$q_t = \phi_0 + \phi_1 P_t + e_t$$

contain v_t, u_t

$$Cov(P_t, e_t) \neq 0$$



Idea

1) regress X on Z (instrumental variables) and any other exogenous X's.

$$X = \underbrace{\phi_0 + \phi_1 Z}_{\text{exogenous variation in } X} + \underbrace{v}_{\text{endogenous variation in } X}$$

"Two Stage Least Squares"

take $\hat{X} = \hat{\phi}_0 + \hat{\phi}_1 Z$

2) regress Y on $\hat{X}$ $\quad y = \beta_0 + \beta_1 \hat{X} + e$

If $Cov(Z, e) = 0$ (exclusion restriction)

$Cov(Z, X) \neq 0$ (relevance)

then $\hat{\beta}_1^{IV}$ is unbiased

Back to supply + demand example

$$q_{st} = \beta_0 + \beta_1 p_t + \beta_2 Z_{st} + u_t$$

$$q_{dt} = \alpha_0 + - \alpha_1 p_t + \alpha_2 Z_{dt} + V_t$$

$$q_{dt} = \alpha_0 + -\alpha_1 \underset{\uparrow \atop z_{st}}{p_t} + \alpha_2 z_{dt} + v_t$$

instruments in demand equation z_{dt}, z_{st} ← price instrument in demand equation

" " supply equation z_{dt}, z_{st}

↗ price instrument ↖ instrument for itself

Focus on demand curve :

$$q_{dt} = \underline{x}_t^T \cdot \underline{\alpha} + v_t = \begin{bmatrix} 1 & p_t & z_{dt} \end{bmatrix} \begin{bmatrix} \alpha_0 \\ -\alpha_1 \\ \alpha_2 \end{bmatrix} + v_t$$

$$\underset{T\times 1}{\underline{q}_d} = \underset{T\times 3}{\tilde{X}} \cdot \underset{3\times 1}{\underline{\alpha}} + \underset{T\times 1}{\underline{v}}$$

run OLS : $\underset{3\times 1}{\hat{\underline{\alpha}}} = \underbrace{\left(\underset{3\times T}{\tilde{X}^T} \cdot \underset{T\times 3}{\tilde{X}}\right)^{-1} \underset{3\times T}{\tilde{X}^T} \cdot \underset{T\times 1}{\underline{q}_d}}$ = matrix version of $\dfrac{Cov(\tilde{X}, q_d)}{Var(\tilde{X})}$

for IV or 2SLS :

1) regress $\tilde{X}$ on $\underset{\uparrow \atop 1,\, z_{st},\, z_{dt}}{\tilde{Z}}$ $\qquad \underbrace{\left(\underset{3\times T}{\tilde{Z}^T} \underset{T\times 3}{\tilde{Z}}\right)^{-1} \underset{3\times T}{\tilde{Z}^T} \underset{T\times 3}{\tilde{X}}}_{3\times 3}$

coefficient matrix in first stage

predict $\hat{\tilde{X}}$ by multiplying by these coefficients

$$\underset{T\times 3}{\hat{\tilde{X}}} = \underbrace{\underset{T\times 3}{\tilde{Z}} \cdot \underset{3\times 3}{\left[(\tilde{Z}^T \tilde{Z})^{-1} \cdot \tilde{Z}^T \cdot \tilde{X}\right]}}$$

2) regress q_d on $\hat{\tilde{X}}$

$$\underset{3\times 1}{\hat{\underline{\alpha}}_{2SLS}} = \left(\hat{\tilde{X}}^T \cdot \hat{\tilde{X}}\right)^{-1} \hat{\tilde{X}}^T q_d$$

$$= \underline{\left[\tilde{X}^T \tilde{Z} \cancel{(\tilde{Z}^T \tilde{Z})^{-1}} \cancel{\tilde{Z}^T \tilde{Z}} (\tilde{Z}^T \tilde{Z})^{-1} \tilde{Z}^T \tilde{X}\right]^{-1} \tilde{X}^T \tilde{Z} (\tilde{Z}^T \tilde{Z})^{-1} \tilde{Z}^T q_d}$$

$$= \left[\tilde{X}^T \tilde{Z} (\tilde{Z}' \tilde{Z})^{-1} \tilde{Z}^T \tilde{X}\right]^{-1} \tilde{X}^T \tilde{Z} (\tilde{Z}^T \tilde{Z})^{-1} \tilde{Z}^T q_d$$

In special case where have only 1 instrument for each endogenous variable, can simplify to

In special case where have only 1 instrument for each endogenous variable, can simplify to

$$\hat{\alpha}_{IV} = (\underset{\sim}{Z}^T \underset{\sim}{X})^{-1} \underset{\sim}{Z}^T q_d$$

matrix version of

$$\frac{Cov(q_d, \underset{\sim}{Z})}{Cov(\underset{\sim}{X}, \underset{\sim}{Z})}$$

relevance

"weak IV problem"

$\downarrow$

$Cov(X, Z)$ is low

$\rightarrow$ unstable or biased coefficients

(Generalized) Method of Moments

$$y_t = \beta_0 + \beta_1 x_{1t} + \beta_2 x_{2t} + \varepsilon_t$$

x_{2t} is endogenous

z_{2t} a good IV

$$\Rightarrow E(\varepsilon_t) = \frac{1}{T}\sum_{t=1}^{T}(y_t - \beta_0 - \beta_1 x_{1t} - \beta_2 x_{2t}) 1 = 0$$

$$\Rightarrow E(\varepsilon_t x_{1t}) = \frac{1}{T}\sum_{t=1}^{T}(y_t - \beta_0 - \beta_1 x_{1t} - \beta_2 x_{2t}) x_{1t} = 0$$

3 eq + 3 unknowns $\beta_0, \beta_1, \beta_2$

$$\cancel{E(\varepsilon_t x_{2t}) = \frac{1}{T}\sum_{t=1}^{T}(y_t - \beta_0 - \beta_1 x_{1t} - \beta_2 x_{2t}) x_{2t} = 0}$$

replace with

$$\Rightarrow E(\varepsilon_t z_{2t}) = \frac{1}{T}\sum_{t=1}^{T}(y_t - \beta_0 - \beta_1 x_{1t} - \beta_2 x_{2t}) z_{2t} = 0$$

$$\cancel{0 \cdot \beta_0 + 0 \cdot \beta_1 + 0 \cdot \beta_2} = 0 \cdot z_{2t}$$

$1, x_{1t}, z_{2t}$ are $\underline{Z}_t$ vector

rewriting system of equations

$$\underset{3\times T}{\underset{\sim}{Z}^T} \; \underset{T\times 1}{\varepsilon} = \underset{3\times 1}{\underline{0}}$$

$$\underset{\sim}{Z}^T \cdot (y - \underset{\sim}{X} \hat{\beta}_{IV}) = \underline{0}$$

$$\underset{\sim}{Z}^T y = \underset{\sim}{Z}^T \underset{\sim}{X} \cdot \hat{\beta}_{IV}$$

$$(\underset{\sim}{Z}^T \underset{\sim}{X})^{-1} \underset{\sim}{Z}^T y = \hat{\beta}_{IV}$$

$$(\underset{\sim}{Z}^T \underset{\sim}{X})^{-1} \underset{\sim}{Z}^T \underset{\sim}{y} = \hat{\beta}_{IV}$$

Extra instruments then I can add moments to the system

parameter vector

$$\underline{m}^{GMM}(\underline{\theta}) = \begin{bmatrix} \frac{1}{T}\sum_{t=0}^{T} \varepsilon_t \\ \frac{1}{T}\sum_{t=0}^{T} x_{1t}\varepsilon_t \\ \frac{1}{T}\sum_{t=0}^{T} z_{2t}\varepsilon_t \\ \frac{1}{T}\sum_{t=0}^{T} z_{3t}\varepsilon_t \\ \frac{1}{T}\sum_{t=0}^{T} z_{4t}\varepsilon_t \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

5 eq., 3 unknowns

use first 3,

use 1, 2, 4

use 1, 2, 5

"average the estimates"

optimal $\underset{\sim}{W} = \underset{\sim}{\Omega}^{-1}$

$$\min_{\underline{\theta}} \underbrace{\underline{m}^{GMM}(\underline{\theta})^T}_{1\times 5} \cdot \underbrace{\underset{\sim}{W}}_{5\times 5} \cdot \underbrace{\underline{m}^{GMM}(\underline{\theta})}_{5\times 1}$$

$(\beta_0, \beta_1, \beta_2)$

$$\underset{\sim}{\Omega} = E\left[\underbrace{\underline{m}(\underline{\theta})}_{5\times 1} \cdot \underbrace{\underline{m}(\underline{\theta})^T}_{1\times 5}\right]$$

expected square of moments (variance of moments)

one-step GMM $\rightarrow$ $\underset{\sim}{\Omega}(\underline{\theta})$ into objective function + optimize

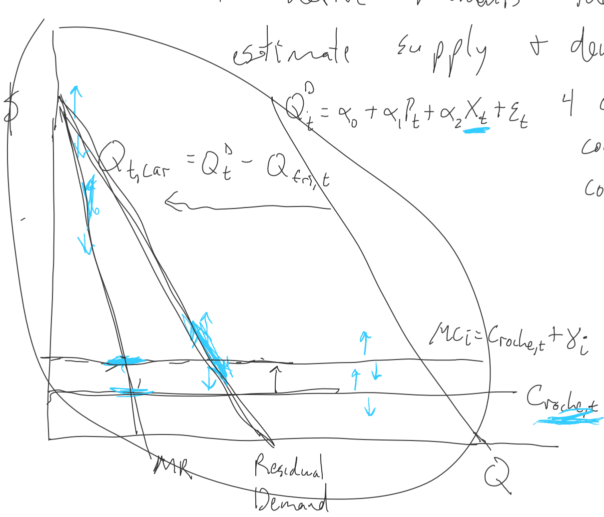
two-step GMM: ① use $\underset{\sim}{I}$ for $\underset{\sim}{W}$. estimate $\tilde{\underline{\theta}}$

② $\min\ \underline{m}(\underline{\theta})^T \left(\underset{\sim}{\Omega}(\tilde{\underline{\theta}})\right)^{-1} \underline{m}(\underline{\theta})$

iterative GMM $\rightarrow$ two-step over + over until $\underline{\theta}$ estimates converge to something

Vitamin Cartel Paper

– use the theory oligopoly (Cournot) and cartels to derive moments that are needed to estimate supply + demand, strategic behavior



4 cournot oligopolists globally

competitive fringe $\rightarrow$ price takers

Cournot $\rightarrow$ subtract the competitive

① fringe supply from aggregate deman

② Derive "residual demand"

③ Optimize over residual demand

Demand side:

Population of large demand countries

$$Q_t^D = \underline{\alpha}_0 + \underline{\alpha}_1 P_t + \underline{\alpha}_2 X_t + \varepsilon_t$$

data $\underline{Q_{fri,t}} = k_t$ or

$= \lambda_t P_t$ or

$= k + \lambda_t P_t$

3 cases

Supply : First Order Condition in a Cournot Problem

$$P_t + \frac{dP}{dQ_t} \cdot q_{it} = c_{i,t} \quad \text{for firm } i = 1, \ldots, 4$$

$$MR = MC$$

$$= c_{roche,t} + \gamma_i + \underset{\text{residual}}{\eta_{it}}$$

Eqm : $Q_t^D = Q_{fri,t} + Q_{car,t}$

Demand, Supply, Eqm give us the moments:

Moments :

Demand : $\underline{m}_1(\underline{\theta}) = \sum_y \left(Q_y - \alpha_0 - \alpha_1 \overline{P}_y - \alpha_2 \overline{X}_y \right) \cdot \begin{bmatrix} 1 \\ \overline{X}_y \\ c_{roche,y} \end{bmatrix} = 0$

3 equations
3 unknowns

annual data

monthly data averaged to annual

Supply : $\underline{m}_{2,i}(\underline{\theta}) = \sum_y \left(\overline{P}_y - \frac{dP}{dQ} \cdot q_{i,y} - c_{roche,y} - \gamma_i \right) \begin{bmatrix} 1 \\ \overline{X}_y \end{bmatrix}$

Eqm :
averaged across firms

$$\underline{m}_3(\underline{\theta}) = \sum_t \left(P_t + \frac{1}{n} \frac{dP}{dQ} \left(\alpha_0 + \alpha_1 P_t + \alpha_2 X_t - Q_{fri,t} \right) - c_{roche,t} - \frac{1}{n} \sum_{i=1}^{n} \gamma_i \right) \cdot X_t$$

$$\min_{\substack{\alpha_0, \alpha_1, \alpha_2 \\ \gamma_1, \gamma_2, \gamma_3, \gamma_4}} \begin{bmatrix} \underline{m}_1^T & \underline{m}_2^T & \underline{m}_3^T \end{bmatrix} \underset{\sim}{I} \begin{bmatrix} \underline{m}_1 \\ \underline{m}_2 \\ \underline{m}_3 \end{bmatrix}$$

IV: remaining ideas

→ bias in 2SLS, relevance
machine learning in "first stage"

SVAR as IV's

then ARCH + GARCH

simple: $\hat{\beta}_{IV} = \frac{Cov(Z,Y)}{Cov(Z,X)}$ $\qquad \hat{\beta}_{OLS} = \frac{Cov(X,Y)}{Var(X)}$

First stage F-stat:

$y_t = \beta_0 + \beta_1 x_t + \beta_2 w_{1t} + \dots + \beta_k w_{k-1,t} + \varepsilon_t$ $\qquad E(x_t \varepsilon_t) \neq 0$

w's are exogenous

first stage:

$x_t = \alpha_0 + \alpha_1 z_{1t} + \alpha_2 z_{2t} + \dots + \alpha_p z_{pt} + \alpha_{p+1} w_{1t} + \dots + \alpha_{p+k+1} w_{k-1,t} + u_t$

First stage F is F-stat from an F-test

of $H_0: \alpha_0 = \alpha_1 = \alpha_2 = \dots = \alpha_p = 0$

Not a great idea to just throw in everything

if some Z's are weak predictors of X

F-stat $\downarrow$

If I've got a lot of candidates for Z,

maybe a machine learning, "regularization" exercise can help

e.g. LASSO in first stage

run regression $x_t = \alpha_0 + \alpha_1 z_{1t} + \alpha_2 z_{2t} + \dots + \alpha_p z_{pt} + \overbrace{\beta w_t}^{\text{other controls}} + u_t$

$$\min_{\alpha's,\beta} \underbrace{\sum_{t=0}^{T} u_t^2}_{OLS} + \underbrace{\lambda \sum_{i=0}^{p} |\alpha_i|}_{\text{regularization penalty}} \qquad \lambda \in (0,1)$$

tuning parameter

choose λ by 10-fold cross validation

- split sample into 10 equally sized chunks
- estimate model ~~#2~~ on first 9, use coefficients

  to predict x_t in 10^{th} chunk

  calculate $\hat{x}_t - x_t$ to get prediction error

repeat for each of 10 chunks to get an ~~average~~

mean squared prediction error

for a given λ!

repeat whole process for different λ's

- pick λ with the lowest mean squared prediction error.

Once we have a λ we like,

run LASSO in first stage using that λ.

Once we have a λ we like,

run LASSO in first stage using <u>that</u> $\underline{\lambda}$.

→ that will keep some Z's and drop other Z's

Last step: use the retained Z's in 2SLS.