Econ 312

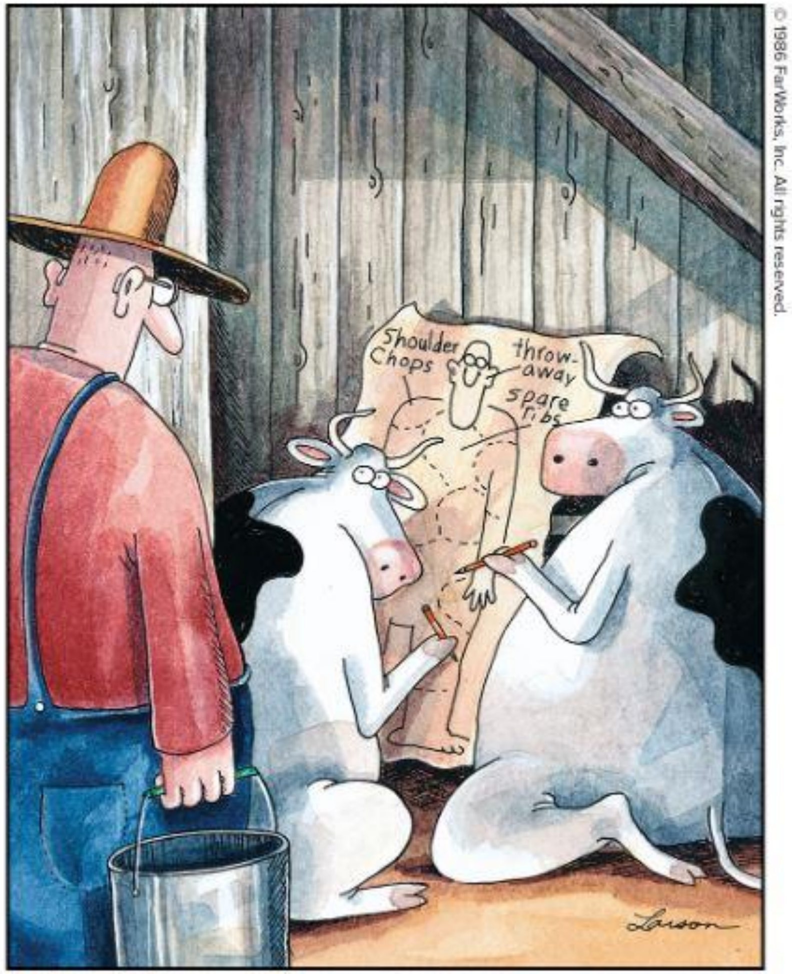
Wednesday, April 17

Identification in Simultaneous Equations

Readings: Wooldridge, Chapter 16

Class notes: 142 - 149

Today’s Far Side offering



Farmer Brown froze in his tracks; the cows stared wide-eyed back at him. Somewhere, off in the distance, a dog barked.

Hmmmm

Context and overview

- We have introduced the instrumental-variables estimator for equations in which there are endogenous regressors
- We now begin a discussion of estimation of systems of equations
- The focus today will be conditions that allow the identification of structural parameters in our models
- We will examine this in the context of an extended example of a supply-demand system with price and quantity endogenous

System estimation vs. single-equation estimation

- Suppose we have an equation in which one of the regressors is endogenous
- What is the model that determines the endogenous regressor?
  - Need that in single-equation IV to choose instruments
- Are we interested in the parameters of the "other equation" that determines the endogenous regressor?
  - If so, then we should estimate both equations
  - If not, then it's usually fine to use IV on the single equation of interest
- System estimation involves estimating all of the related equations

The identification problem

- In VARs, we had to make a set of assumptions in order to identify the structural shocks and do impulse-response functions
- Exactly the same problem arises in instrumental-variables estimation
    - We must assume that our instruments are uncorrelated with u and that the do not directly affect y
    - These are our identifying assumptions or restrictions that allow us to do instrumental variables and estimate structural parameters
- We examine this problem through an extended example of a two-equation demand-supply system

Model I: No exogenous variables

$$\text{Demand curve: } Q = \alpha_0 + \alpha_P P + u$$

$$\text{Supply curve: } Q = \beta_0 + \beta_P P + v$$

- Set Q equal across the two equations and solve for P (See notes)

$$P = \frac{\alpha_0 - \beta_0}{\beta_P - \alpha_P} + \frac{u - v}{\beta_P - \alpha_P} \equiv \pi_{P0} + \varepsilon_P$$

- Plug this value of P back into one equation and solve for Q

$$Q = \frac{\beta_P \alpha_0 - \alpha_P \beta_0}{\beta_P - \alpha_P} + \frac{\beta_P u - \alpha_P v}{\beta_P - \alpha_P} \equiv \pi_{Q0} + \varepsilon_Q$$

- The parameters are the coefficients of the reduced-form equations
- We can estimate these by OLS

Identification in Model I

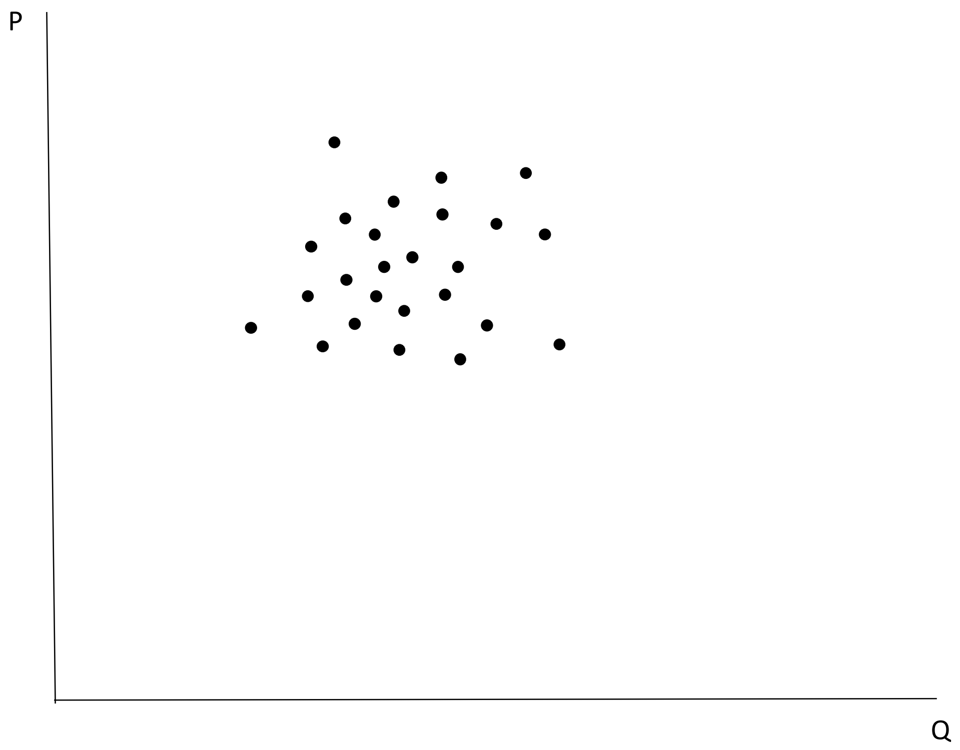
- Structural parameters wanted (4): α_0, α_P, β_0, β_P
- Coefficients we can estimate (2): $\pi_{P,0}$, $\pi_{Q,0}$

- No way to find 4 structural parameters from only 2 reduced-form parameters
  - We can't find *any* of the structural parameters here
  - Nothing is identified: neither demand nor supply equation

Graphical illustration

Demand curve: $Q = \alpha_0 + \alpha_P P + u$

Supply curve: $Q = \beta_0 + \beta_P P + v$



- How do we get a supply or demand curve from the scatter of points?
  - We cannot
  - All of the variation is due to u and v
  - There is no way to identify which points are demand shifts and which are supply shifts
- Neither the demand curve nor the supply curve is identified

Model II: Income (M) affects demand

$$\text{Demand curve: } Q = \alpha_0 + \alpha_P P + \alpha_M M + u$$

$$\text{Supply curve: } Q = \beta_0 + \beta_P P + v$$

- Solving for reduced form as before

$$P = \frac{\alpha_0 - \beta_0}{\beta_P - \alpha_P} + \frac{\alpha_M}{\beta_P - \alpha_P} M + \frac{u - v}{\beta_P - \alpha_P} \equiv \pi_{P0} + \pi_{PM} M + \varepsilon_P$$

$$Q = \frac{\beta_P \alpha_0 - \alpha_P \beta_0}{\beta_P - \alpha_P} + \frac{\alpha_M \beta_P}{\beta_P - \alpha_P} M + \frac{\beta_P u - \alpha_P v}{\beta_P - \alpha_P} \equiv \pi_{Q0} + \pi_{QM} M + \varepsilon_Q$$

- If we estimate the π parameters by OLS we have 4 coefficients
- There are now 5 structural parameters, so identification cannot be complete

Identification in Model II

- Structural parameters wanted (5): α_0, α_P, α_M, β_0, β_P
- Coefficients we can estimate (4): $\pi_{P,0}$, $\pi_{Q,0}$, $\pi_{P,M}$, $\pi_{Q,M}$

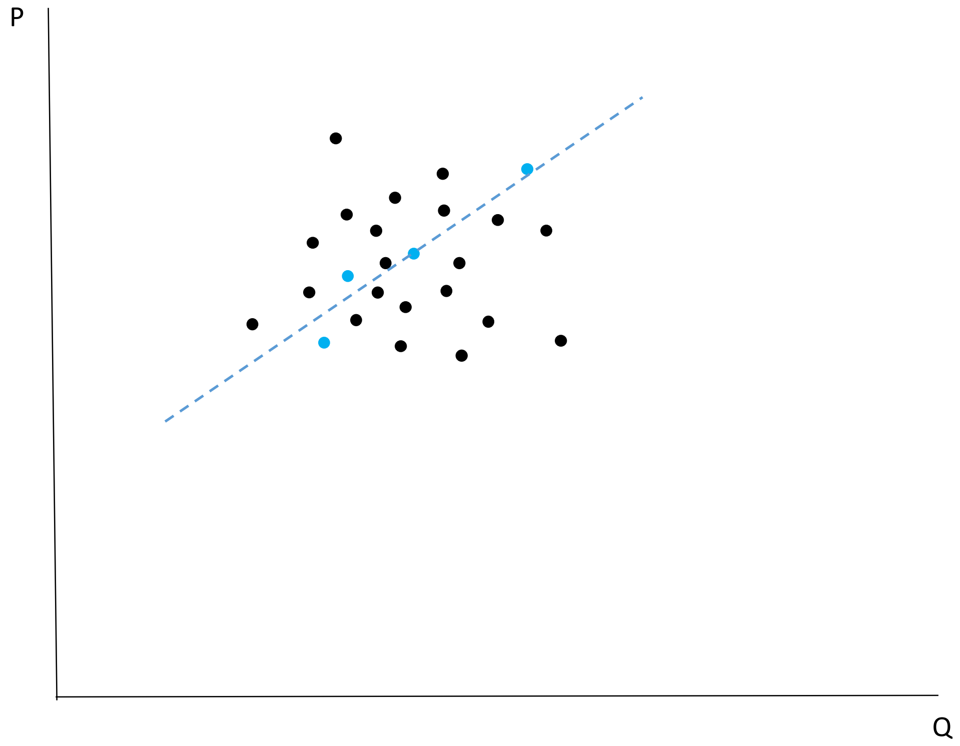
$$\frac{\pi_{QM}}{\pi_{PM}} = \frac{\dfrac{\alpha_M \beta_P}{\beta_P - \alpha_P}}{\dfrac{\alpha_M}{\beta_P - \alpha_P}} = \beta_P$$

- β_P in the supply equation can be identified! Likewise β_0 (in notes)
- Supply equation is identified by presence of exogenous variable in demand equation. Demand equation is still not identified

Graphical illustration

Demand curve: $Q = \alpha_0 + \alpha_P P + \alpha_M M + u$

Supply curve: $Q = \beta_0 + \beta_P P + v$



- Suppose that the blue points have different levels of income
- We know that income affects demand *but not supply*
- We can use the variation in income to represent shifts in the demand curve that identify a supply curve (the blue line)
- We have no variable that shifts (only) the supply curve, so cannot identify demand curve

Model III: Income (M) affects both curves

$$\text{Demand curve: } Q = \alpha_0 + \alpha_P P + \alpha_M M + u$$

$$\text{Supply curve: } Q = \beta_0 + \beta_P P + \beta_M M + v$$

- Solving for reduced form,

$$P = \frac{\alpha_0 - \beta_0}{\beta_P - \alpha_P} + \frac{\alpha_M - \beta_M}{\beta_P - \alpha_P} M + \frac{u - v}{\beta_P - \alpha_P} \equiv \pi_{P0} + \pi_{PM} M + \varepsilon_P$$

$$Q = \frac{\beta_P \alpha_0 - \alpha_P \beta_0}{\beta_P - \alpha_P} + \frac{\alpha_M \beta_P - \beta_M \alpha_P}{\beta_P - \alpha_P} M + \frac{\beta_P u - \alpha_P v}{\beta_P - \alpha_P} \equiv \pi_{Q0} + \pi_{QM} M + \varepsilon_Q .$$

- If we estimate the π parameters by OLS we have 4 coefficients
- There are 6 structural parameters, so we again do not have enough information to identify structural parameters uniquely

Identification in Model III

- Structural parameters wanted (6): α_0, α_P, α_M, β_0, β_P, β_M
- Coefficients we can estimate (4): $\pi_{P,0}$, $\pi_{Q,0}$, $\pi_{P,M}$, $\pi_{Q,M}$

- No longer possible to identify any parameters
- Because *M* affects BOTH curves, we cannot use variations in *M* to trace out the supply curve
- The crucial identifying restriction that worked to identify the supply curve in Model II:
  - *M* affected demand BUT NOT SUPPLY
  - Now that *M* affects both curves, we cannot use it for identification

Model IV: *M* affects demand, *R* affects supply

Demand curve: $Q = \alpha_0 + \alpha_P P + \alpha_M M + u$

Supply curve: $Q = \beta_0 + \beta_P P + \beta_R R + v$

- Solving for reduced form yields

$$P = \frac{\alpha_0 - \beta_0}{\beta_P - \alpha_P} + \frac{\alpha_M}{\beta_P - \alpha_P} M - \frac{\beta_R}{\beta_P - \alpha_P} R + \frac{u - v}{\beta_P - \alpha_P} \equiv \pi_{P0} + \pi_{PM} M + \pi_{PR} R + \varepsilon_P$$

$$Q = \frac{\beta_P \alpha_0 - \alpha_P \beta_0}{\beta_P - \alpha_P} + \frac{\alpha_M \beta_P}{\beta_P - \alpha_P} M - \frac{\beta_R \alpha_P}{\beta_P - \alpha_P} R + \frac{\beta_P u - \alpha_P v}{\beta_P - \alpha_P} \equiv \pi_{Q0} + \pi_{QM} M + \pi_{QR} R + \varepsilon_Q$$

- If we estimate the π parameters by OLS we have 6 coefficients
- There are now 6 structural parameters, so identification is at least possible

Identification in Model IV

- Structural parameters wanted (6): α_0, α_P, α_M, β_0, β_P, β_R
- Coefficients we can estimate (6): $\pi_{P,0}$, $\pi_{Q,0}$, $\pi_{P,M}$, $\pi_{Q,M}$, $\pi_{P,R}$, $\pi_{Q,R}$

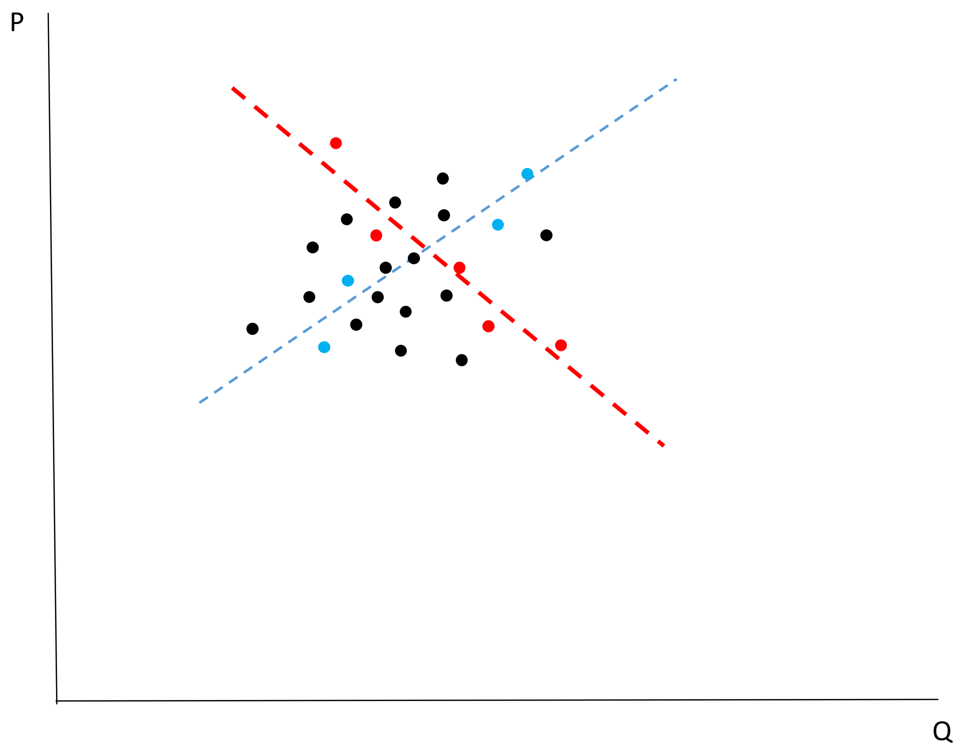
$$\frac{\pi_{QM}}{\pi_{PM}} = \frac{\dfrac{\alpha_M \beta_P}{\beta_P - \alpha_P}}{\dfrac{\alpha_M}{\beta_P - \alpha_P}} = \beta_P \qquad \frac{\pi_{QR}}{\pi_{PR}} = \frac{\dfrac{\alpha_P \beta_R}{\beta_P - \alpha_P}}{\dfrac{\beta_R}{\beta_P - \alpha_P}} = \alpha_P \qquad \begin{aligned} \pi_{PM}\left(\beta_P - \alpha_P\right) &= \alpha_M \\ -\pi_{RM}\left(\beta_P - \alpha_P\right) &= \beta_R \end{aligned}$$

- All of the parameters in BOTH equations can be identified
  - Supply equation is identified by *M* in demand equation
  - Demand equation is identified by *R* in supply equation

Graphical illustration

Demand curve: $Q = \alpha_0 + \alpha_P P + \alpha_M M + u$

Supply curve: $Q = \beta_0 + \beta_P P + \beta_R R + v$



- Blue points have different levels of income, which affects demand but not supply
- Red points have different levels of rainfall, which affects supply but not demand
- Variation in income identifies a supply curve (blue line)
- Variation in rainfall identifies a demand curve (red line)

Model V: *M* affects demand, *R* and *W* affect supply

Demand curve: $Q = \alpha_0 + \alpha_P P + \alpha_M M + u$

Supply curve: $Q = \beta_0 + \beta_P P + \beta_R R + \beta_W W + v$

- Solving for reduced form yields

$$P = \frac{\alpha_0 - \beta_0}{\beta_P - \alpha_P} + \frac{\alpha_M}{\beta_P - \alpha_P} M - \frac{\beta_R}{\beta_P - \alpha_P} R - \frac{\beta_W}{\beta_P - \alpha_P} + \frac{u - v}{\beta_P - \alpha_P} \equiv \pi_{P0} + \pi_{PM} M + \pi_{PR} R + \pi_{PW} W + \varepsilon_P$$

$$Q = \frac{\beta_P \alpha_0 - \alpha_P \beta_0}{\beta_P - \alpha_P} + \frac{\alpha_M \beta_P}{\beta_P - \alpha_P} M - \frac{\beta_R \alpha_P}{\beta_P - \alpha_P} R - \frac{\beta_W \alpha_P}{\beta_P - \alpha_P} W + \frac{\beta_P u - \alpha_P v}{\beta_P - \alpha_P} \equiv \pi_{Q0} + \pi_{QM} M + \pi_{QR} R + \pi_{QW} W + \varepsilon_Q.$$

- If we estimate the π parameters by OLS we have 8 coefficients
- There are now 7 structural parameters, so identification is at least possible and we might have EXTRA information (overidentification)

Identification in Model IV

- Structural parameters wanted (7): α_0, α_P, α_M, β_0, β_P, β_R, β_W
- Coefficients we can estimate (8): $\pi_{P,0}$, $\pi_{Q,0}$, $\pi_{P,M}$, $\pi_{Q,M}$, $\pi_{P,R}$, $\pi_{Q,R}$, $\pi_{P,R}$, $\pi_{Q,R}$
- We can do the same thing as Model IV to identify all parameters, but now

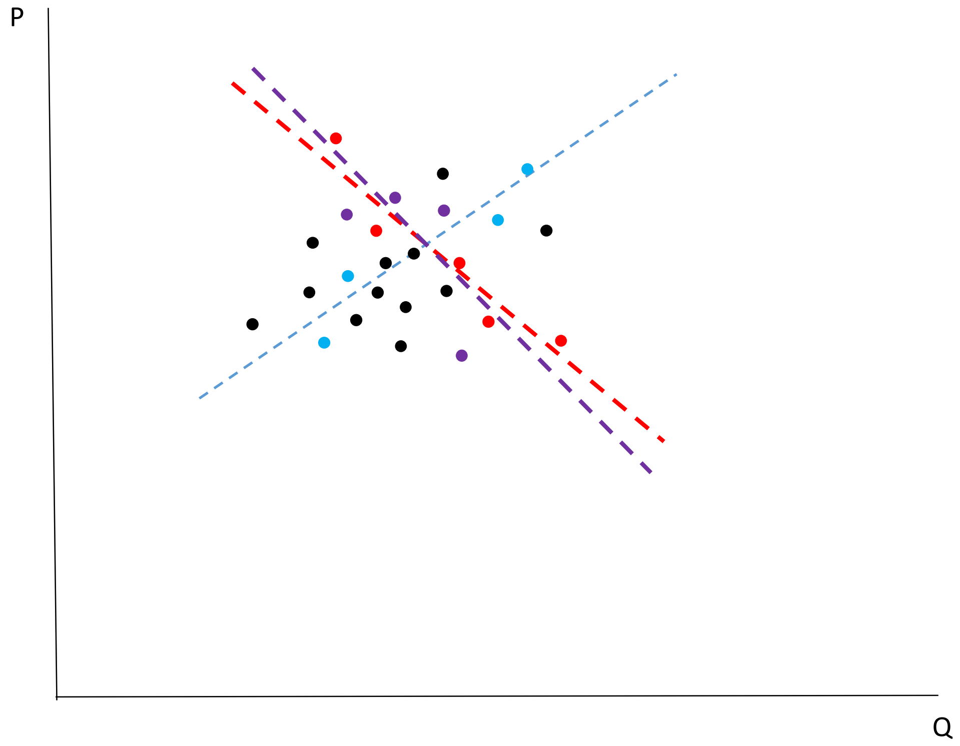
$$\alpha_P = \frac{\pi_{QR}}{\pi_{PR}} \quad \text{OR} \quad \alpha_P = \frac{\pi_{QW}}{\pi_{PW}}$$

- With TWO variables shifting the supply equation (and not in demand), we have two different choices of how to identify
- Model is overidentified
- Do we get the same answer? Is $\dfrac{\pi_{QR}}{\pi_{PR}} = \dfrac{\pi_{QW}}{\pi_{PW}}$?
  - Testable overidentifying restriction

Graphical illustration

Demand curve: $Q = \alpha_0 + \alpha_P P + \alpha_M M + u$

Supply curve: $Q = \beta_0 + \beta_P P + \beta_R R + \beta_W W + v$



- Blue points have different levels of income; affects D but not S
- Red points have different levels of rainfall; affects S but not D
- Purple points have different wage; affects S but not D
- Variation in income identifies a supply curve (blue line)
- Variation in rainfall OR variation in wage identifies a demand curve (red line or purple line)

Basic principles of identification

- For each endogenous regressor in ONE equation, you need at least one exogenous variable that is in the OTHER equation but not the first
  - This exogenous variable is an instrument, just like in IV
  - For the system to be identified, each equation must be identified
- If we have extra exogenous variables in the other equation, then the first equation is overidentified
  - Exactly the same as having excess instruments in IV
  - We can test the overidentifying restrictions to assess model validity

Review and summary

- In this lecture, we have considered the identification of structural coefficients in systems of equations
- We can only estimate reduced forms (using OLS)
- If an equation is identified, then its structural parameters can be obtained as functions of the (estimable) reduced-form coefficients
- It is possible for only some of the equations in a system to be identified
- If there are multiple alternative ways of estimating the structural parameters of an equation from the reduce-form coefficients, then the equation is overidentified; we can test the validity of the model by testing overidentifying restrictions

Something different

If these were normal times, I'd be inviting you to come and see our marimba band perform on Saturday at our teacher/leader's annual concert.

This year, the best I can do is offer a clip from last year's performance, recorded on our old camcorder with crappy sound (even worse when compressed here!) and people walking in front of the camera.

What’s next?

- In this class, we learned the details of identification in systems of simultaneous regression equations
- Our next class (April 20) will consider how to estimate entire systems of equations together rather than separately
- We will learn the method of seemingly unrelated regressions for systems with no endogenous regressors
- We will then combine this method with 2SLS to get the three-stage least-squares estimator for systems of equations