Structural Estimation and that of Anderson and van Wincoop (AvW) Model

Anton Yang

Lecture developed for graduate international trade curriculum
International Trade Theory AgEc 644 (Purdue University)

April 6, 2021

Structural Estimation (SE)

Structural estimation, SE, or sometimes used interchangeably with structural econometrics:

- What is it?
- Why is it relevant to us?
- How is it relevant?

What is it?

- Structural estimation is a technique for estimating **deep** "structural" parameters of theoretical economic models. The term is inherited from the **simultaneous equations model** (Wikipedia).
- Koopmans (1965) (*Cowles Commission Paper*):

  Statistical inference, from observations to economic behavior parameters, can be made in two steps: inference from the observations to the parameters of the assumed joint distribution of the observations, and inference from that distribution to the parameters of the structural equations describing **economic behavior**.

- Economists use similar advanced techniques as computer and chemical engineers do (e.g., Ferris *et al.*(2009)).
- But the phrase "**structural estimation**" $\longrightarrow$ Economists.

SE and "Machine" Learning

Sound tangent, but highly relevant to us as economists:

- Both are rapidly emerging scientific literatures.
- Scientists in the areas of econometrics and computer science seem to be still learning from each other.
- There are analogies, contrasts and synergies between the two.
- Would AI replace the role of human creativity (SE techniques) and knowledge in model structures and inferences?
- Iskhakov *et al.*(2020), Igami (2020), Keane and Neal (2020)

Some deeper questions to ask in the future:

- How much of the world do we want to predict?
- How much of the world do we want to explain?

Why is it relevant?

- When we run a linear regression of Y on X, we essentially try to understand the relationship between X and Y based on some **empirical observations** (data!):

$$Y = X\beta + \epsilon$$

- which can be *loosely* generalized from empirical gravity model (the so-called naive version of economic gravity!)

$$\ln(\text{Trade Flows}_{ij}) = \alpha_i \ln(GDP_i) + \alpha_j \ln(GDP_j) - \gamma \ln(Distance_{ij}) + \epsilon_{ij}$$

- This, however, is not based on economic "theory".
- This does not explain much the world.
- Thus economists think hard to incorporate principles of economic behavior using **structural equations**.

How is it relevant?

- The empirical form of **observational model** does not draw inference from economic behavior, so what?
- The structural model has solid economic theoretical foundations which have been built up over many decades.
- They are made convenient for **counterfactual analysis** with supply, demand and many more economic behaviors.
- For example, we care about how firm behaves, corresponding to price shifts due to exogenous shocks.
- The firm's behavioral elasticities are not only determined by prices, but also by other conditions, such as resources.

Simple Math Representation (if any)

Loosely speaking, we say that there is a k x k matrix represented by Γ that categorizes the relationships between economic "outputs" $Y = [y_1, y_2, ...y_k]$, then

$$Y\Gamma = X\alpha + U$$

is a **structural form** (given relevant assumptions hold), and if we multiply by Γ^{-1} on both sides, then

$$Y = X\beta + \epsilon$$

is the original **reduced-form**,

where $\beta = \alpha\Gamma^{-1}$ and $U = \epsilon\Gamma^{-1}$.

- This says that OLS is a special version of SE.

The actual problems are a little more complicated than this!

"Deep" Structural Parameters

- The parameters that are being estimated and identified from structural equations are called **structural parameters**.
- Also called "**policy-invariant**" parameters.
- As opposed to **reduced-form** parameters.
- See Lucas Critique, i.e., Lucas (1976).

Existing critiques on structural parameters:

- How deep are the deep parameters? Altissimo *et al.* (2002)
- How structural are the structural parameters? Fernández-Villaverde *et al.* (2007)

Perhaps, at least in DSGE models, we need the data to tell us some *portions* of the parameter *drifting* and *stability*.

Structural Parameters in CGE and NQT

- The merit of "deepness" is that economists can implement economic counterfactual analysis (and in succinct ways!).
- Without at least some **fixed** parameters being **identified**, it is impossible to make counterfactual predictions.
- This is true in both CGE and NQT literature.
- This is true in both global and local counterfactuals.
- See Allen and Arkolakis (2015) lecture 9.
- The **existence** and **uniqueness** in levels equilibrium imply the same for counterfactual equilibrium which are calculated using structural parameters.
- In GTAP, these counterfactuals are in percent change form (written in GEMPACK), but parameters typically come from elsewhere, they are "deep" but not quite "structural".

Structural Equations and Equilibrium System

From optimization point of view, these are just **"constraints"**.

- The difference is, statisticians and computer scientists often try to minimize the residuals using **data structure**, whereas economists do more or less so using **economic structure**.
- These are not just for trade economists, but for economists.
- See Su and Judd (2012), *Econometrica*.
- ...and six years ago, Balistreri and Hillberry (2006).
- These are pioneers who use an MPEC approach to tackle **mixed complementarity problems (MCP)** and (usually highly) non-linear system of structural equations.

Complementarities as Constraints

Recall that for a canonical LP program

$$\begin{aligned} \min_{x} \quad & c^T x \\ \text{s.t.:} \quad & Ax \geq b, \quad x \geq 0. \end{aligned}$$

- The complementarity relationships written in "$\perp$" are:

$$\begin{aligned} 0 \leq c - A^T \lambda \quad & \perp \quad x \geq 0, \\ 0 \leq Ax - b \quad & \perp \quad \lambda \geq 0. \end{aligned}$$

  (Necessary and sufficient **optimality conditions**)

- The product of the non-zero element(s) on each side of "$\perp$" is zero, or has an either or relationship.

- See also Choi (Computational Economics, 2015), p.308.

Complementarities in Economics

- E.g., in some homothetic CES-gravity, it can be shown that

$$\left(\sum_i \beta_i p_i^{1-\sigma}\right)^{\frac{1}{1-\sigma}} \geq PU \quad \perp \quad u \geq 0$$

$$PU = \frac{E}{u}$$

- It implies that *production of utility* must be zero whenever strict inequality constraints holds in equilibrium (familiar?).
- That is, there is no utility produced, if the price is zero.
- The equality is the unit expenditure defining equation.

Complementarities and MPEC in GE

Structural estimation is sort of combining some of the math in pg. 11 and economics in pg. 12 (previous slide).

- In structural estimation, we formulate economic relationship as constraints to the objective function, such that
- necessary and sufficient optimality conditions are met.
- This makes it an MPEC problem (this case: **MCP** + **NLP**),
- where MCP belongs to the family of variational inequality.
- A key advantage of MPEC is that it allows structural parameters and objectives to be fully theory-consistent.
- MCPs in economic optimization problems that involve in market-clearing is implicitly a **bi-level** problem.

Complementarities and MPEC in GE

- The choice of one economic agent depends on quantities chosen by other agents in a market and multiple optimizers.
- Moreover, supply objective (subject to constraints) solves demand objective (subject to constraints) in GE.
- This means GE systems are essentially bi-level problems.
- This means reduced-form cannot fully identify a GE problem.
- MPEC can evaluate several structural parameters **simultaneously** while holding **GE consistent**, where
- a operational system of equations must be in order: n equations are associated with n unknowns.

AvW Model and BH Critiques

- AvW's contribution is huge in trade literature, but
- previous slide leads to our discussion of a major critic made by Balistreri and Hillberry (06/07/08) on AvW Model.
- BH points out that a fully operational GE model of AvW Model of aggregate trade should be $4n$ equations, not $2n$.
- AvW substituted equilibrium scaled prices ($\beta_i p_i$) into consumer price index and solve them for demand equations.
- **Substitutions** that produce $2n$ equation system generate a numeraire shift in the counterfactuals, not GE consistent.
- This means that β's are not structurally estimated while assuming border costs are symmetric.

AvW Model and BH Critiques

- That's another reason why MPEC-style estimator is useful.
- It allows for a full econometric calibration of the model.
- Information on β's and other "deep" structural parameters are extracted from the data variations directly.
- We may impose asymmetric border costs as constraints. (could be a good coding question)
- MPEC fosters a transparent link to the counterfactuals.

Re-estimating AvW using $4n$ GE System

(1) Income definition:

$$Y_i = FOB_i E_i^0.$$

(2) Goods market-clearing condition:

$$E_i^0 = \sum_j \left[\frac{Y_j}{FOB_i} \left(\frac{\alpha_i FOB_i t_{ij}}{P_j} \right)^{1-\sigma} \right].$$

(3) Unit expenditure function (**CES price index**):

$$P_j = \left[\sum_i (\alpha_i FOB_i t_{ij})^{1-\sigma} \right]^{1/(1-\sigma)}.$$

(4) Income balance:

$$U_i P_i = Y_i.$$

Choosing an Objective Function

- Choosing an objective function in SE is an art.
- In structural econometrics, we may set up a variety of different estimators in the objective function, such as Least Squares, MLE, GMM and PPML, like in reduced-form approach.
- It is important to ask which one is a better candidate.
- Current trade literature likes to use PPML.
- At a minimum, due to some zero trade flows in the data.
- See also Fally (2015), *Journal of International Economics*.
- We stay simple here, and use **least squares** optimization.

Least Squares (LS) Optimization

- It is almost always a good idea to start out with LS.
- It is more difficult to mess up in the coding.
- Choose this as an operational and "numerical" baseline.
- In BH (06/07), we set up the objective as follows

$$\min \quad \sum_i \sum_j [z_{ij} - \hat{z}_{ij}]^2$$

- where $\hat{z}_{ij}$ is the fitted import share of i in j.
- Location j buys everything from the world (incl. j).
- Denominator is then the income (=expenditure).
- See BH (07) Equation (8).

Normalization

- **Normalization** is critical in structural estimation. (standardization/regularization in statistics)
- This is another art, but often very difficult.
- Uniqueness is up to a **scale**.
- w/o establishing a scale, problems that are not solvable in OLS are not identifiable in structural econometric approach.
- See Yang and Preckel (2020), equation (3.7), in it

$$\beta_k u^{e_k(1-\alpha_k)}$$

- is not structurally identifiable without normalization.

Normalization

- To see this, $u = \rho\nu^{\delta}$, for some $\rho > 0$ and $\delta > 0$
- ν will be as good as u for explaining the data.
- But if we substitute u with $\rho\nu^{\delta}$.
- We'll get

$$\beta_k \rho^{e_k(1-\alpha_k)} \nu^{\delta e_k(1-\alpha_k)},$$

- which generalizes to

$$\tilde{\beta}_k \nu^{\tilde{e}_k(1-\alpha_k)}.$$

The **scales** (scaling factors) are not themselves identifiable from the data, so the parameters are not fully identified.

Normalization

In this problem, BH (07) establishes the scale of utility that is consistent with AvW:

$$(5) \quad P_{\text{Alabama}}^{1-\sigma} = \sum_i \left[\frac{Y_i}{\sum_j Y_j} \left(\frac{t_{\text{Alabama},j}}{P_i} \right)^{1-\sigma} . \right]$$

This normalization constraint then adds to (1) - (4).

- How to test if we have the right normalization?
- Operationally, this can be challenging.
- In GAMS, we can implement the test of **initial guess**. (at very low cost!)
- Unchanged solutions or "optimum" with alternative starting values does not guarantee the right normalization.
- But when this happens, however, it would almost unambiguously guarantee that the normalization is **incorrect**.

Identification and Parameter Robustness

- In SE, we want to ask whether there are additional degrees of freedom in the parameter spaces that are not identified.
- Yang and Preckel (2020) find a structural method which can be called "post-hoc" latitude test of parameter robustness.
- This method is simple, and does not require extensive additions to the computer codes.
- The theory is intuitive, they ask: whether there is room for moving the model parameters such that the primitive objective is improved?
- Can be useful in SE due to high linearities and complexities.

Numeric Stability

- Econometricians often scale the data for numeric stability.
- This implementation should not impact the model due to the magnitude of the data and that proportions are unchanged.

```
53  *       Rescale the nominal data to $billions for better numeric stability
54  Scalar  scale  /1e-3/;
55  Y0(i)=Y0(i)*scale;
56  ship(i,j)=ship(i,j)*scale;
```

- Good data scaling helps approaching a feasible solution.
- Bad scaling might lead to occasional software failures.
- This together works with the art of choosing **starting values**.
- In this case, both GDPs and trade flows are scaled by 0.001.

Extensions and the Future

- Most existing SE literature adopts what can be referred to as **hyperparameter** in machine learning literature.
- For instance, many of them say $\sigma = 5$.
- Hillberry and Yang (2020) essentially say that σ can be identified using economic theoretical structure alone.
- This potentially improves the interpretability of SE.
- But should we let/trust the data (to) do the job?
- On the other hand, emerging ML literature continues to improve the interpretability of model components after the model is well-learned.
- What will happen in the foreseeable future?