Advanced Applied Econometrics
Week 1 - Intro and Static Discrete Choice

Peter Haan

21 April 2023

Introduction

- **EMAILS: mail list for the course**
- Organization, see outline

If anything **unclear**, please ask.

Presentation: name, study status, main econ interest

Course schedule

1. Methods
2. Paper discussion $\Rightarrow$ you **need to read at home**
3. Practical session (numerical methods)
   $\Rightarrow$ you **need** to do problem sets
   $\Rightarrow$ you **need** STATA and Python
   $\Rightarrow$ you **need** to practice coding

Evaluation

For students who need to be evaluated

(1) **Special problem sets**

- individually or in pairs
- provide commented code, results, documentation, tables

(2) **Exam**

- 21.7.2023;

Plan for today

- Structural econometrics
- Examples of a structural model
- Practical session,
- Python and example with Max Blesch

What is structural econometrics?

What is structural econometrics?

What is structural econometrics?

Structural econometric models...

- "...combine explicit **economic theories with statistical models**" (Reiss and Wolak, 2007)
- "...[infer] underlying parameters that represent **tastes, technology, and other primitives** from observations of individual households and firms..." (Adda and Cooper, 2003)
- ... estimate **features of a data generating process** (i.e., a model) that are (assumed to be) invariant to the policy changes or other counterfactuals of interest (Haile, 2022)

Descriptive vs structural analyses

- Descriptive analysis: establish facts about observable quantities
- RCT
- Policy evaluation
- Economic model

Program Evaluation

Haile (2022):

- Program evaluation (indeed, any type of so-called "causal inference" is always a form of structural estimation. It requires a set of maintained hypotheses about the world (i.e., a model) allowing one to define and identify a counterfactual quantity of interest.
- TT, ATE, LATE, QTE, etc. are all precisely defined only under a well specified model of how the data are being generated. Any suggestion that these objects are **"model free" is nonsense**.
- Causality is always defined by a counterfactual.

Reduced form vs structural

Haile (2022):

- "Reduced form" is sometimes used to mean "equation I won't derive, justify, or take questions on, but which I will nonetheless treat as causal when I talk about conclusions. **This is just bad science.**"

Other views and definitions by Rust (2010), thoughtful (and spiteful) comments on Keane (2013) in Fritjers (2013) vs Rust (2013).

An Introductory Example

How much revenue raised by **introducing income taxes**?

Simple answers?

Effects of tax change I

How much **additional revenue (AR)**

1. AR=new tax rate * wage * number of hours worked
2. AR=new tax rate * wage * **hours worked post tax**

How many hours worked post tax?

Effects of tax change II

How many hours worked post tax?

- estimate **observed corr between wages and hours**
- **interpret as L^S-elasticity**?

Effects of tax change III

using theory... why might **number of hours worked change**?

1. **hours margin**: change in hours for (pre- & post-) workers
2. **participation margin**: change in participation pre vs. post-tax:

Effects of tax change IV

3. **life-cycle** margin:
   changing taxes may influence optimal allocation of work over life
   (eg. if progressivity is increased...)

4. **wage** margin: **equilibrium (gross) wage may change**

Conclude: Predicting reform effect (AR) requires structural model. But: Model choice depends on **margin** and **identification**.

Before the practical session

10 minute break

Practical Session - Economic model simple version of van Soest 1995

Question: The Flat Party wins elections and wants to **introduce a Flat Income Tax (FIT)**.

We want to know **how much revenue is raised under FIT**.

Why do we need a structural model? Alternatives?
How to model this ?

Practical Session

What minimal necessary ingredients for a model?

1. decision to participate
2. decision over hours worked (discretize - why?)
3. earnings of participants (pre- and post reform)
4. NB. recognize heterogeneity (e.g. due to different non-labour incomes)

(1) Participation and hours

How do we model this?

(1) Participation and hours II

Individuals **maximize utility** from consumption & leisure

$$U(c,h) = \gamma \left[\frac{c^\theta}{\theta} - \alpha\, h\right] + \varepsilon_h$$

by **choosing hours per week** $h \in [0, 10, 20, 30, 40]$
faced with constraint posed by wage.

(2) Consumption

Simplest model of consumption in this context?

(2) Consumption II

$$c = w * h$$

What does this imply?
Simplest models of wages?

(3) Wages

Assume workers draw wage rates from normal wage offer distribution *before* choosing hours

$$\log w = \mu_w + \epsilon_w$$

where ε_w from $N[0, \sigma_w]$

Estimation and Identification

Use maximum likelihood to estimate...

1. parameters of the wage distribution
2. taste parameters of leisure-consumption choice

How is the model identified? How can we estimate a causal effect based on the model?

Coding

Principles

1. Use a **script** and **document your code**. Always.
2. There is **always** a solution. Not one.
3. One solution is **fast**,
4. another is **easy to read** & understand,
5. another is **flexible**.
6. You won't find any immediately. No matter. **Try again.**

Simulate to estimate

Good practice: Simulate before estimating. Why?

In practice: Simulation and Estimation

1. Simulate economic behaviour in model with your favourite parameters.
2. Generate simulated data.
3. Save the data. Forget the parameters.
4. Write the likelihood function as a function of data and parameters.
5. Maximize likelihood of observing your simulated data w.r.t. parameters.
6. Rediscover your favourite parameters.