Lecture 1: Introduction

ResEcon 703: Topics in Advanced Econometrics

Matt Woerman
University of Massachusetts Amherst

Agenda

Today

- Course overview
- What is "structural estimation?"
- Getting started with R

Upcoming

- For next time
  - Reiss and Wolack (2007), Sections 1–4
  - Install R software

Course Overview

Course Website

`github.com/woerman/ResEcon703`

I will use this GitHub repository to post lecture slides, R code, problem sets, datasets, etc.

My Info

Matt Woerman

- Email: `mwoerman@umass.edu`
- Office: 218 Stockbridge Hall
- Office hours: Tuesday, 2:00–3:00 pm

Best way to communicate with me

- Short public question: Ask in class
- Short private question: Email with [ResEcon 703] in the subject
- Longer question: Come to office hours

About Me

- I study energy and environmental economics, industrial organization, and applied econometrics
  - Market power in wholesale electricity markets
  - Role of electricity in agricultural groundwater extraction
  - Design of carbon markets and other environmental policies
  - Tools for designing field experiments using panel data
- This is my first year as a professor and first time teaching this course
  - You can play a role in shaping the design of this course, for yourself and for future classes!
- My wife is a professor in the Biology Department at UMass
  - "Dr. Woerman" is not a unique identifier, so call me "Matt"
- Pronouns: he/him/his

About You

Introduce yourself

- Name
- Pronouns
- Department
- Research interests
- Anything else you want us to know?

Course Description

You've already taken

- ResEcon 701: Probability Theory and Statistical Inference
- ResEcon 702: Econometric Methods
  - Classical linear regression model
  - "Treatment effect" estimation

(If you have not taken ResEcon 702, please see me to determine if this course is appropriate for you)

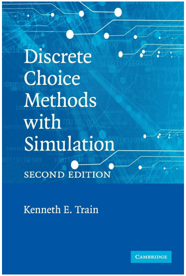
Isn't that enough? What else is there?

- Nonlinear regression models
- Structural estimation
- Discrete choice models

Course Goals

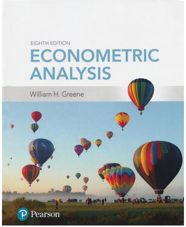
1. Gain an in-depth understanding of the most common structural estimation methods in modern empirical economics
   - Maximum likelihood estimation
   - Nonlinear least squares
   - Generalized method of moments
2. Develop the technical ability to apply these methods to your own research
3. Apply these methods to discrete choice models motivated by the random utility model
   - Logit model
   - Generalized extreme value models (nested logit)
   - Mixed logit model (random coefficients logit model)

Textbooks



Discrete Choice Methods with Simulation (Second Edition) by Kenneth E. Train

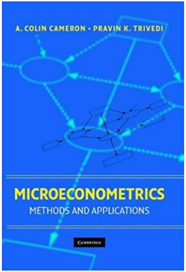
- Available for free at: `eml.berkeley.edu/books/choice2.html`
- Paperback copy is only $40



Econometric Analysis (Eighth Edition) by William H. Greene

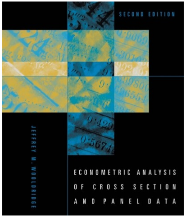
- Same textbook you used in ResEcon 702
- Seventh Edition is also fine

Other References



Microeconometrics: Methods and Applications by A. Colin Cameron and Pravin K. Trivedi

- More applied than many econometrics texts



Econometric Analysis of Cross Section and Panel Data (Second Edition) by Jeffrey M. Wooldridge

- More advanced material than other textbooks

Readings

We will read papers to see how these structural estimation methods have been applied to

- Energy economics
- Environmental economics
- Experimental economics
- Health economics
- Industrial organization
- Labor economics
- Public economics

Software

We will use the R statistical programming language in this course

But I already know Stata/Matlab/Python/SAS/Julia. Why R?

- R is free and open source
- R is powerful and flexible
  - Basic statistics, data cleaning, linear regression, matrix algebra, simulation methods, structural estimation, data visualization, etc.
- R is favored by employers

How can I learn R?

- R tutorial next class
- Many R resources available for free
- First problem set will be a (relatively) gentle introduction to R

You do not have to use R. But I will not provide any support or partial credit for work done in other programming languages.

Grades

Your final grade will be made up of

- Problem sets: 4 at 15% each (60% total)
- Final exam: 30%
- Attendance and participation: 10%

Problem Sets

Problems sets will simulate the kind of analysis you will do when conducting your own research

- Apply the estimation methods you learn in class
- Interpret your results
- Draw policy-relevant conclusions

Rules for problem sets

- You can work in groups of up to three people (I recommend you do)
- Each person must submit a unique set of answers
- You must submit your code with your write up
- You can only use “canned” routines when told to do so

See syllabus for tentative problem set schedule

Final Exam

Final exam will be similar to problem sets

- Take-home, not in class
- Estimation, interpretation, etc.
- At least a week to complete (actual exam timeline TBD)

How the final exam differs from problem sets

- Will require roughly twice the effort of a problem set
- SINGLE-AUTHORED! No collaboration, consultation, etc.

More details to come toward the end of the semester

Attendance and Participation

Attendance is required

- You can miss at most two lectures and still receive full points

"Participation" is required

- Come to class prepared
- Complete the reading before class
- Bring your laptop to work through examples together

See syllabus for tentative reading schedule

What Is “Structural Estimation?”

Structural Econometric Model

Definition according to Reiss and Wolack (2007)

- Framework that "combines explicit economic theories with statistical models"

Economic theory

- Tells us how a set of observed endogenous variables (y) are related to a set of observed exogenous variables (x)
- May also relate the y variables to unobserved variables (ξ)
- Specifies a function form ($g(\cdot)$) and unknown parameters (Θ)

$$y = g(x, \xi, \Theta)$$

Statistical assumptions

- Give a joint distribution of x and ξ

$$\ell(y, x \mid \Theta) \quad \text{and} \quad E(y \mid x, \Theta)$$

Auction Example

We observe the winning bid (y) and the number of bidders (x) from many auctions, and we want to understand the relationship between the number of bidders and the winning bid

Non-structural ("reduced-form") approach

- Regress winning bids on number of bidders
- No economic theory, institutional details, microeconomic fundamentals, etc.
- Is this a causal estimate? No!

Structural approach

- Incorporate economic and institutional details into relationship
  - Combine auction theory with statistical assumptions to estimate underlying distribution of valuations, risk preferences, and differences between auction formats
- Gets closer to causality in this setting
- Required to simulate counterfactuals

Structural vs. Non-Structural Models

Is a structural model always better than a non-structural model?

- NO! The right approach depends on your research question, data, institutional details, etc.

Non-structural ("reduced-form") models

- With a good research design, non-structural models can provide
  - Causal estimates
  - Less reliance on researcher assumptions
  - Transparent assumptions, estimation, and results
- Without a good research design, interpretation is less clear

Structural models

- Estimate unobserved microeconomic fundamentals ("structural parameters")
- Allow for counterfactual of policy simulations
- Can be used to compare competing economic theories

Constructing a Structural Econometric Model, Step 1

Start with economic theory

- Description of the economic setting
  - Markets, institutions, agents, information
- List of primitives
  - Technologies, preferences, endowments
- Exogenous variables
  - Constraints, regulations, shifters
- Objective function and decision variables
  - Utility maximization and quantities demanded, profit maximization and input quantities
- Equilibrium concept
  - Walrasian equilibrium with price-taking, Nash equilibrium with quantity selection

Constructing a Structural Econometric Model, Step 2

Transform economic model into econometric model

- Unobservables that account for the data not perfectly fitting the economic model
  - Researcher uncertainty about the economic setting
  - Agent uncertainty about the economic setting
  - Optimization error by agents
  - Measurement error in observed variables

Constructing a Structural Econometric Model, Step 3

Estimate the econometric model

- Functional forms
- Distribution assumptions
- Estimation method
- Specification tests

Example of a Structural Model

We want to estimate the elasticities of a Cobb-Douglas production function when we observe output (Y_{it}), capital (K_{it}), and labor (L_{it})

$$Y_{it} = A_i K_{it}^{\alpha} L_{it}^{\beta} \eta_{it}$$

1. Re-write the production function

$$\ln(Y_{it}) = \ln(A_i) + \alpha \ln(K_{it}) + \beta \ln(L_{it}) + \ln(\eta_{it})$$

2. Define

$$\gamma_i = \ln(A_i) \quad \text{and} \quad \varepsilon_{it} = \ln(\eta_{it})$$

3. Make statistical assumptions

$$\varepsilon_{it} \sim N(0, \sigma^2) \quad \text{and} \quad E(\varepsilon_{it} \mid K_{it}, L_{it}) = 0$$

Then we can use OLS to estimate

$$\ln(Y_{it}) = \gamma_i + \alpha \ln(K_{it}) + \beta \ln(L_{it}) + \varepsilon_{it}$$

A More Complex Example of a Structural Model

We observe the bids $(b_1, \ldots, b_N)$ from procurement auctions with N risk-neutral bidder, and we want to estimate the underlying joint density of costs $(f(c_1, \ldots, c_N))$

1. Economic theory tells us each firm will maximize expected profit

$$E[\pi_i(b_1, \ldots, b_N)] = (b_i - c_i) \Pr(b_i < b_j \; \forall j \neq i \mid c_i)$$

2. Taking the derivative gives the first-order condition

$$b_i = c_i - \Pr(b_i < b_j \; \forall j \neq i \mid c_i) \left(\frac{\partial \Pr(b_i \leq b_j \; \forall j \neq i)}{\partial b_i} \right)^{-1}$$

3. Assume all firms know $f(c_1, \ldots, c_N)$

$$b_i = c_i + \frac{\int_{c_i}^{\infty} [1 - F(\tau)]^{N-1} d\tau}{[1 - F(c_i)]^{N-1}}$$

A More Complex Example, Continued

$$b_i = c_i + \frac{\int_{c_i}^{\infty} [1 - F(\tau)]^{N-1} d\tau}{[1 - F(c_i)]^{N-1}}$$

$$\vdots$$

99. Substituting in additional assumptions

$$c_i = b_i - \frac{1}{N-1} \left(\frac{1 - G(b_i)}{g(b_i)} \right)$$

To estimate $f(c_1, \ldots, c_N)$

1. Nonparametrically estimate $G(\cdot)$ and $g(\cdot)$ from data on bids
2. Estimate c_i using this estimated distribution and density of bids
3. Nonparametrically estimate $f(c_1, \ldots, c_N)$ from estimated costs

Structural Estimation

Some structural models can be estimated using OLS or related regression

- Easy and fast to implement
- Estimation procedure and underlying assumptions are transparent
- Results are easily interpreted

Some structural models require more advanced estimation methods

- Structural model cannot be simplified to a linear regression model
- Methods are broadly defined as "structural estimation"

This course will focus on "structural estimation" that follows from this second class of structural models

Getting Started with R

Installing R

Installing R is *usually* straightforward

- Download (`cran.r-project.org`) and install R
- Download (`www.rstudio.com/products/rstudio/download`) and install RStudio Desktop (Open Source License)

What is the difference between R and RStudio?

R is like a car's engine. It is the program that powers your data analysis.

RStudio is like a car's dashboard. It is the program you interact with to harness the power of your "engine."

Fundamentals of R

Everything is an object

```r
foo
```

Every object has a name and value

```r
foo <- 2
```

You use functions on objects

```r
mean(foo)
```

Functions come in packages/libraries

```r
library(mlogit)
```

Playing Around in R

Basic math operations

```
1 + 2
## [1] 3

1 + 2 == 3
## [1] TRUE

(1 + 2) / 3
## [1] 1

2^3
## [1] 8
```

Playing Around in R

Basic math with objects

```
a <- 1
b <- 2
a + b
## [1] 3

c <- a + b
c
## [1] 3

(a + b) / c
## [1] 1

b^c
## [1] 8
```

Playing Around in R

Functions

```r
exp(2)
## [1] 7.389056

sqrt(3)
## [1] 1.732051

d <- c(1, 2, 3, a, b, c)
d
## [1] 1 2 3 1 2 3

mean(d)
## [1] 2

max(d)
## [1] 3

min(d)
## [1] 1
```

Playing Around in R

Matrix math

```
mat_a <- matrix(c(1, 2, 3, 4), nrow = 2)
mat_a
##      [,1] [,2]
## [1,]    1    3
## [2,]    2    4

2 * mat_a
##      [,1] [,2]
## [1,]    2    6
## [2,]    4    8

mat_a %*% mat_a
##      [,1] [,2]
## [1,]    7   15
## [2,]   10   22
```

Playing Around in R

Install and load packages

```r
install.packages(c('tidyverse', 'mlogit', 'gmm'))
library(tidyverse)
library(mlogit)
library(gmm)
```

Announcements

For next time

- Reiss and Wolack (2007), Sections 1–4
- Install R software