Midterm Exam Study Guide

ECON 6343: Structural Econometrics

Exam Date: Tuesday, November 11, 2025

Duration: 75 minutes

Format: Closed everything (no notes, books, devices, AI)

General Advice

- Read each question carefully
- Show your work for partial credit
- Circle or box final answers for calculation problems
- Bring a basic scientific calculator
- Budget your time appropriately across questions

Topic Areas to Review

1. Distributional Properties and Foundations

Key Concepts:

- Understand the properties of Type I Extreme Value distributions
- Know how the logistic distribution relates to Type I EV distributions
- Be able to work with moments (means and variances) of these distributions
- Review how to manipulate expectations and variances of random variables
- Understand independence assumptions and how they affect variance calculations

Practice:

- Derive moments of transformed random variables
- Practice proofs involving expectations and variances
- Review properties of commonly-used distributions in discrete choice models

---

2. Identification Fundamentals

Key Concepts:

- Distinguish between identification, estimation, and inference
- Understand what it means for a parameter to be "identified"
- Know that identification is a population concept, not a sample concept
- Recognize that sample size and statistical significance are irrelevant to identification
- Understand the types of assumptions required for identification:
  - Statistical assumptions (e.g., sufficient variation)
  - Behavioral assumptions (e.g., exogeneity, exclusion restrictions)

Common Misconceptions:

- Large sample ≠ identification
- Statistical significance ≠ identification
- Tight confidence intervals ≠ identification

Practice:

- Be able to explain what identification means in plain language
- Be prepared to critique flawed reasoning about identification
- Review identification arguments from papers you've read

---

3. Model Validation and Testing

Key Concepts:

- Understand the purpose of validation tests across different empirical approaches
- Recognize that validation tests build confidence but cannot prove untestable assumptions
- Know the parallel between validation in experimental and structural approaches
- Understand what assumptions are untestable in each approach
- Know common validation strategies for:
  - Randomized experiments (balance checks)
  - Difference-in-differences (parallel trends)
  - Structural models (model fit, out-of-sample prediction)

Important Points:

- Balance checks don't prove randomization worked (only check observables)

- Pre-trends don't guarantee post-treatment parallel trends
- Model fit doesn't prove behavioral assumptions are correct
- All approaches require some leap of faith

Practice:

- Compare and contrast validation approaches across methods
- Be able to identify untestable assumptions in different research designs
- Think about what validation tests can and cannot tell us

---

4. Discrete Choice Model Specification

Key Concepts:

- Understand how discrete choice model coefficients are scaled
- Know why logit coefficients cannot be directly compared across specifications
- Understand the role of error variance normalization in identification
- Recognize differences between:
  - Comparing coefficients in OLS vs. logit
  - Comparing marginal effects vs. coefficients
  - Comparing parameters across nested vs. non-nested models

Key Distinction:

- OLS: Coefficients are marginal effects (interpretable)
- Logit: Coefficients are scaled by error variance (not directly interpretable)

Practice:

- Review the latent variable framework for binary choice models
- Understand why adding variables changes all coefficients in logit
- Practice calculating and interpreting marginal effects

---

5. Reproducible Research Practices

Key Concepts:

- Understand the difference between absolute and relative file paths

- Know principles of portable, reproducible code
- Recognize code that violates portability principles
- Understand project organization and directory structures
- Know why relative paths facilitate collaboration and version control

Bad Practices:

- Hard-coded absolute paths (e.g., `C:/Users/YourName/...`)
- Machine-specific file locations
- Code that only works on one computer

Good Practices:

- Relative paths from project directory
- Portable code that runs on any machine
- Clear project organization

Practice:

- Review how to navigate directories with relative paths (`../`, `./`)
- Practice converting absolute paths to relative paths
- Think about project workflow and file organization

---

6. Dynamic Discrete Choice Models

Key Concepts:

- Understand the $\mathbb{E}\max$ operator and its role
- Know why $\mathbb{E}\max$ appears in dynamic programming
- Understand what the expectation and maximum are taken over
- Review the Bellman equation structure
- Understand forward-looking behavior and continuation values
- Know that $\mathbb{E}[\max] \neq \max[\mathbb{E}]$ (Jensen's inequality)

Key Questions:

- What does the $\max$ operator represent? (optimization over choices)
- What does the $\mathbb{E}$ operator represent? (expectations over uncertainty)

- Why can't we separate them?

Practice:

- Review Rust (1987) and the bus engine replacement model
- Work through the Bellman equation step by step
- Understand how agents make forward-looking decisions

---

7. Extensions of Multinomial Logit

Key Concepts:

- Know the fundamental limitation of multinomial logit: IIA property
- Understand the Red Bus/Blue Bus paradox
- Know how nested logit addresses IIA:
  - Grouping alternatives into nests
  - Allowing correlation within nests
  - The role of the nesting parameter λ
- Know how mixed logit addresses IIA:
  - Random coefficients across individuals
  - Flexible correlation patterns
  - Can approximate any discrete choice model
- Understand trade-offs between nested logit and mixed logit

Classic Example: Red Bus/Blue Bus problem illustrates IIA violation and motivates flexible models

Practice:

- Review Train Chapters 4 (GEV) and 6 (Mixed logit)
- Work through examples of substitution patterns
- Understand how different models generate different predictions

---

Study Strategy

1. **Review lecture slides** for topics covered through Class 22 (Learning models)
2. **Re-read key sections of Train:**

- Chapters 1-2 (Basics)
- Chapter 3 (Logit)
- Chapter 4 (GEV/Nested Logit)
- Chapter 6 (Mixed Logit)

3. **Review problem sets** PS1-PS6, focusing on conceptual understanding
4. **Practice derivations** involving expectations, variances, and distributions
5. **Review clean code principles** from early lectures
6. **Understand identification arguments** from papers discussed in class

---

Topics NOT Covered on This Exam

Based on the syllabus, the following topics come after the midterm and will NOT be tested:

- Constrained optimization and equilibrium models (Class 24)
- Type II Extreme Value models (Class 25)
- Relative correlation restrictions (Class 26)
- Machine learning (Classes 27-28)
- Specific content from your presentations/referee reports

Formula Sheet

You may want to memorize or be familiar with:

- Mean of standard Type I EV: $\gamma \approx 0.5772$
- Variance of standard Type I EV: $\frac{\pi^2}{6}$
- Variance of logistic distribution: $\frac{\pi^2}{3}$
- Properties of expectations and variances for independent random variables
- Basic discrete choice probability formulas

Final Tips

- **Don't just memorize formulas** – understand the concepts and intuition
- **Practice explaining concepts in words** – many questions ask for explanations

- **Review the "why" not just the "what"** – understand motivation behind methods
- **Budget time** – don't spend too long on any one question
- **Get good sleep** the night before!

Good luck! 🍀