

Professor John Gliebe  
Portland State University  
Toulan School of Urban Studies and Planning

Lectures: Cramer Hall #1  
Phone: 503-725-4016  
E-mail: gliebej@pdx.edu

## USP 510 Discrete Choice Modeling

Description: This course presents the theory and practice underlying the formulation and estimation of models of individual discrete choice behavior with applications to travel, travel related and other choices. The course will provide students with an understanding of the theory, methods, application and interpretation of multinomial logit (MNL), nested logit and other members of the Generalized Extreme Value (GEV) family of models. It will also include an introduction to mixed logit models.

### Course Objectives:

- Develop an understanding of the theory of choice modeling.
- Familiarization with and understanding of the art, science and methods of discrete choice modeling.
- Develop the ability to use theory, judgment and statistical analysis to obtain enhanced model specifications.
- Become familiar with relevant software and its use.
- Become familiar with current research in choice modeling.

Format: Classes will be conducted in lecture/discussion format. Readings for each topic will be assigned and expected to be read before class. Homework assignments will provide an opportunity to experiment with and learn about model specification and testing. Unless otherwise specified, *all homework will be submitted as a single electronic document with figures, tables and attachments included in the main document using the report guidelines specified in Appendix A of the syllabus.* A paper copy should also be submitted. Homework will be discussed during the class immediately following their due date. Unless prior arrangements are made with me, homework submitted after results are discussed in class will be penalized by a letter grade (*e.g.*, an A becomes a B, an A- becomes a B-, etc.). There is one exception to this rule. Specifically, with prior approval, each student is allowed one extension (of up to 48 hours). Students may collaborate on the approach to and analysis of homework assignments but are expected to prepare and present results and interpretations independently.

Readings: There are three required texts for this course. Students are required to obtain Train (2003). We will also be using Ben-Akiva and Lerman (1985). Course notes and supplemental readings will be distributed via the course web site.

1. The primary text is *A Self Instructing Course in Mode Choice Modeling* (SIC2), Prepared For U.S. Department of Transportation, F.S. Koppelman *et al.*, January 31, 2006. This text will be distributed electronically via the course web site.
2. Students who will be doing research in this area should purchase *Discrete Choice Analysis: Theory and Application to Travel Demand* (DCA) by Ben-Akiva and Lerman (1985 edition), available from The MIT Press, 55 Hayward Street, Cambridge, MA 02142, (1-800-356-0343 or [mitpress-orders@mit.edu](mailto:mitpress-orders@mit.edu)).

3. Students should purchase *Discrete Choice Methods with Simulation* (TRAIN) by Kenneth Train (2003 edition) published by Cambridge University Press and available via Kenneth Train's website at <http://elsa.berkeley.edu/books/choice2.html>. I highly recommend you buy a paperback copy of this book, which may be purchased through many on-line retailers.

Grades: Grades will be based on homework assignments (95%); class participation and attendance (5%). There will be no examinations; however, there may be some short, in-class quizzes to test progress and be sure everyone understands the material. If so, the quiz grades will be rolled into the class participation and attendance score.

Office Hours: My office is in the Urban Center (URB 570Q). I will have formal office hours Thursdays from 1 to 3 p.m. Students may request to meet with me other times as well by e-mailing me for an appointment. It has been my experience that most questions arise when using software to estimate models to homework assignments. As such, it is usually quickest to communicate questions and resolve issues via e-mail to me and/or via a e-mail class.

Acknowledgements: This course, including notes, homework assignments, and data are adapted from those used by Frank S. Koppelman at Northwestern University and Laurie Garrow at Georgia Tech. I am grateful to Frank and Laurie for letting me use this material in developing this course.

Use of Data Beyond Course: The use of data beyond this course is prohibited. Any students desiring to use data for publication purposes must check with the instructor to ensure appropriate permissions to use the data is obtained.

#### Contact Information

For more information contact Prof. Gliebe at [gliebej@pdx.edu](mailto:gliebej@pdx.edu), or 503-725-4016.

#### Required Background

Students are expected to have familiarity with intermediate calculus, intermediate statistics including hypothesis testing, matrix algebra and linear regression.

Planned Topics and Readings

Topics covered will be based on progress and class interests

1. Choice Theory
  - SIC2, Chapters 1-3
  - DCA, Chapter 3
2. Binary and Multinomial Choice Models: Theory & Estimation
  - SIC2, Chapters 4, 5
  - DCA, Chapters 4 (skim) & 5
  - TRAIN, Chapters 2, 3
3. Model Specification Development and Testing
  - SIC2, Chapter 6
4. Nested Logit and Generalized Extreme Value Models
  - SIC2, Chapter 7
  - DCA, Chapter 10 (Except Section 10.7)
5. Sampling and Sample Design
  - DCA, Chapter 8
  - Other materials to be distributed
6. Prediction with Disaggregate Choice Models
  - SIC2, Chapter 8
  - DCA, Chapter 6
7. Models of Spatial Choice
  - Materials to be distributed
8. Choice Experiments and Stated Preference Models
  - Materials to be distributed
9. Advanced Topics in Discrete Choice Modeling
  - TRAIN, Chapter 4, 6 & 9
  - Other materials to be distributed

## APPENDIX A: GUIDELINES FOR WRITING TECHNICAL REPORTS AND HOMEWORK

Homework assignments provide you with an opportunity to model individuals' transportation choices using actual datasets. Your grades on homework assignments will be based on how well you apply modeling concepts introduced in class and on how well you present your results. The following guidelines should be used when preparing homework assignments. Unless otherwise specified, homework assignments should be submitted to me by e-mail or, preferably through WebCT, along with a paper copy to be handed in during class.

1. Title page including your name, e-mail, and homework assignment number.
2. Introduction. 2-3 lines describing choice situation, dataset used, etc. Be sure to include any assumptions regarding choice set availability restrictions and missing data.
3. Main body. Synthesize modeling steps, key findings, etc.
  - (a) Try to fit tables onto one page and generally following the attached examples. You can use two pages to present models, but make sure the tables can be printed (e.g., tables should not extend beyond margins of the paper). You may also report model comparison statistics in the same or separate tables.
  - (b) Numbers in tables should be rounded to an "appropriate" precision. For example, parameter estimates and t-stats should not extend beyond 3 significant digits.
  - (c) Discussion of results should include sufficient qualitative discussion. Intuitively, do the results make sense (for the dataset/choice situation) in terms of parameter estimates, ratios of coefficients, value of time estimates, etc.?
  - (d) Comparison of models should include both a qualitative discussion and quantitative statistical tests.
  - (e) Feel free to include a *brief* discussion of problems encountered when using software (e.g., convergence problems). Often, knowing what didn't work is valuable feedback for class discussions and future homework assignments.
4. Other requirements
  - (a) Font size should be 12 pt. or higher.
  - (b) Page numbers should be included.
  - (c) All supporting materials (e.g., charts and tables) should be included in a single electronic document. I have found that it is easiest to create tables in Excel, copy them to the clipboard, then paste them into your report document as a "picture-enhanced metafile".

TABLE 1. Example of Model Reporting Format

<b>Multinomial Logit Models</b>	<b>Model 1</b>		<b>Model 2</b>		<b>Model 9</b>	
Log Likelihood (Betas Estimated)	-4071		-3600		-3302	
Log Likelihood (Zero Coefficients)	-7223		-7223		-7223	
Rho-Squared	0.436		0.502		0.543	
Rho-Squared Adjusted	0.436		0.501		0.539	
Number of Observations	5029		5029		5029	
	Parameter t-Statistic		Parameter t-Statistic		Parameter t-Statistic	
<b>Constants</b>						
Drive Alone						
Shared Ride 2	-2.1379	-44.117	-2.3197	-42.158	-2.4561	-35.268
Shared Ride 3+	-3.3045	-40.598	-3.7171	-39.910	-3.7578	-32.060
Transit	-1.9516	-38.576	-1.0126	-11.095	-1.4278	-7.972
Bike	-3.3323	-23.058	-3.0953	-20.011	-2.9672	-8.585
Walk	-1.7735	-19.938	-0.7160	-5.564	-0.4806	-1.605
Total Travel Time (minutes)			-0.0498	-15.394		
In-Vehicle Travel Time					-0.0254	-3.858
Walk Travel Time					-0.0775	-9.807
Bike Travel Time					-0.0330	-2.520
Waiting Time (Transit)						
Waiting Time / Distance (Transit)						
Waiting Time / In(Dist+1) (Transit)					-0.1326	-9.871
Total Cost (cents)			-0.0049	-20.591		
Low Income					-0.0028	-5.145
Medium Income					-0.0028	-7.937
High Income					-0.0021	-5.201
<b>Zero Vehicle Households</b>						
Transit					2.3990	7.574
Bike					1.7378	2.690
Walk					2.9869	6.541
<b>Vehicles Less Than Workers</b>						
Drive Alone						
Shared Ride 2					1.7983	13.842
Shared Ride 3+					1.3388	5.910
Transit					1.8901	11.356
Bike					2.0091	6.036
Walk					1.7640	5.580
<b>Vehicles Equal to Workers</b>						
Drive Alone						
Walk					0.6064	2.295
<b>Work Location in CBD</b>						
Shared Ride 2					0.7109	4.233
Shared Ride 3+					1.2933	5.880
Transit					1.5217	7.377
Walk					1.1825	3.286