

**THE UNIVERSITY OF HONG KONG  
FACULTY OF BUSINESS AND ECONOMICS**

**PhD Course Syllabus: MKTG6006**

**Course Code/Title:** [Course Code] [Empirical Marketing Models]

**Course Description:** This module covers empirical models and structural modelling in marketing and new empirical industrial organization (NEIO) and provides students with deep understanding of data analysis and modelling issue in marketing and NEIO. It includes empirical models on the analysis of scanner panel data at individual or household level as well as aggregate data at store, account, market, regional or national level. The topics include brand choice, category choice (purchase incidence), store choice, purchase quantity, and purchase timing, relating to the various consumer decisions (where to buy, whether to buy, what to buy, and how much to buy), either separately or jointly. It also covers learning, forward-looking behavior, search models, NEIO models (the BLP approach and counterfactual experiments). Estimation methods include MLE, GMM, and SMLE. All topics are empirical in nature. Data and basic Gauss and SAS code are provided for the models covered. Relevant readings are also provided. Students are required to work with raw data, cleaning the data, writing their code, estimating the models, and writing reports.

**Course Objectives:** This course is intended for 1<sup>st</sup> and 2<sup>nd</sup> year Ph.D. students in marketing, economics, and information systems. Students will learn fundamental empirical marketing and NEIO modeling and coding skills, and apply them to real data and research questions.

**Pre-requisite:** None, but students should know statistics, econometrics, economic theories, and have some programming skills.

**Assessment:** 100% coursework; 0% examination

**Remarks:** All PhD courses are non-credit-bearing and will be assessed on a pass/fail basis.

<b>Course Learning Outcomes (CLOs)</b> On completion of this course, students should be able to:	<b>Aligned PLOs*</b>				
	1	2	3	4	5
1. Grasp marketing models and cutting-edge empirical methodologies	X	X	X		
2. Apply cutting-edge marketing models to real research questions		X	X	X	
3. Estimate BLP and run counterfactual experiments		X	X	X	
4. Read and understand most empirical marketing papers			X	X	X
5. Write critical reviewer's report on papers that involve empirical models		X	X	X	X

**\*Programme Learning Outcomes (PLOs) for Research Postgraduate Programme:**

1. Demonstrate critical understanding, at an advanced level, of up-to-date knowledge and research methodology of a particular field
2. Implement effective academic and personal strategies for carrying out research projects independently and ethically
3. Contribute original knowledge in response to issues in their specialist area
4. Communicate research findings at a diverse range of levels and through a variety of media
5. Evaluate one's own research in relation to important and latest issues in the field

## COURSE DETAILS *(subject to change at instructor's discretion)*

**Year/Semester:** 2023-24, Second Semester

**Time/Venue:** Monday: 14:30-17:30, KK1211

**Instructor:** [Prof. Junhong Chu]  
Email: [chu123@hku.hk](mailto:chu123@hku.hk)  
Office: KKL-720 (by appointment)

### I. Teaching and Learning Activities

In-class and Out-of-class Activities <i>(e.g. lectures, class discussion, papers reading, proposal writing)</i>	Expected hour	% of student study effort
1. Lectures	36	20
2. Data analysis and coding	50	40
3. Paper readings	60	15
4. Proposal writing	54	25
Total	200	100%

### II. Assessment

Assessment Components <i>(e.g. assignments, proposal, presentation, examination)</i>	Weight	CLOs to be assessed				
		1	2	3	4	5
1. Assignments	15	X	X	X	X	X
2. Paper replications	40	X	X	X	X	
3. Proposal	30		X	X	X	X
4. Classroom participation	15	X	X	X	X	X
Total	100%					

**Students will be assessed based on the following performance standards:**

Course Grade	Performance Standard
Pass	Come to all classes, finish all assignments, complete two paper replications, and submit and present a proposal
Fail	No submission of any of the two paper replications, or no submission of the proposal.

#### Assessment Component 1 *(Optional)*

Score	Performance Standard

### Assessment Component 2 (Optional)

Score	Performance Standard

### Assessment Component 3 (Optional)

Score	Performance Standard

### Assessment Component 4 (Optional)

Score	Performance Standard

## III. Course Content and Tentative Schedule

[Please provide details here]

#### Class 1 Marketing Data and Quantitative Modeling: An Overview

- Introduction to marketing data sources and data structure (slides)
- Computer demonstration: how to clean data and generate new variables based on IRI consumer scanner panel data

#### Class 2 Logit model with scanner data (1): Model setup and derivatives

- Model setup and derivation
- Computer demonstration: in Excel and Gauss
- Interpretation
- Reading: Guadagni & Little (1983, MS)

#### Class 3 Logit model with scanner data (2): Consumer heterogeneity and model application

- Observed heterogeneity (consumer demographics)
- Unobserved heterogeneity: Latent class model (Discrete distribution)
- Latent class model demonstration in Excel and Gauss
- Logit model applications (slides) on three papers (Hardie, Johnson & Fader 1993, MS; Heilman, Briesch et al 1997; Bowman & Wright, 2000 JMR)
- Reading: Kamakura and Russell (1989, JMR)

#### Class 4 Nested Logit model and Rank order model

- Nested logit model
- Rank order model

#### Class 5 Random coefficients model with scanner data

- Random coefficients model

- Reading: Chintagunta and Vilcassim (1991) – investigating heterogeneity in brand choice models

Class 6 Purchase quantity modelling

- Poisson and NBD models
- Truncated Poisson and NBD models
- Zero-inflated Poisson and NBD models
- Tobit models for limited DV

Class 7 Purchase timing modeling

Class 8 Aggregate logit model (1): Setup and Price Endogeneity

Class 9 Aggregate logit model (2): Customer Heterogeneity (Latent class and continuous distribution)

Class 10 Aggregate logit model (3): Nested Logit and Pricing

Class 11 Structural Modelling and BLP

Class 12 Learning models/Search models

Class 13 proposal presentation

#### IV. Required/Recommended Readings

[Please provide details here]

##### Reference books

1. Train, Kenneth. *Discrete Choice Models with Simulation*, which can be downloaded from <https://eml.berkeley.edu/books/train1201.pdf> There are also Gauss code and problem sets on Train's website
2. Benk-Akiva and Lerman. *Discrete Choice Analysis: Theory and Application to Travel Demand*, MIT press. Kindle version is available.
3. Anderson, de Palma and Thisse. *Discrete Choice Theory of Product Differentiation*, MIT Press.
4. Franses and Paap. *Quantitative Models in Marketing Research*. Cambridge University Press. (A simple read).

##### Relevant readings

##### Discrete choice models on Individual/household scanner data

1. Guadagni and Little. 1983. "A logit model of brand choice calibrated on scanner data", *Marketing Science*, 2(3): 203-238.
2. Swait Joffre and Jordan Louviere. 1993. The role of the scale parameter in the estimation and comparison of multinomial logit models", *JMR*, Aug. 305-14.
3. Pradeep, Chintagunta (1992), "Estimating a Multinomial Probit Model of Brand Choice Using the Method of Simulated Moments," *Marketing Science*, 11 (Autumn), 386-407
4. Kamakura and Russel. 1989. "A Probabilistic Choice Model for Market Segmentation and Elasticity Structure", *Journal of Marketing Research*, vol. 26 (Nov): 379-390. (latent class model to account for consumer heterogeneity)
5. Bruce G. S. Hardie, Eric J. Johnson, and Peter S. Fader. 1993. "Modeling Loss Aversion and Reference Dependence Effects on Brand Choice", *Marketing Science*, 12(4): 378-394.
6. Richard A. Briesch, Lakshman Krishnamurthi, Tridib Mazumdar, and S. P. Raj. 1997. "A Comparative Analysis of Reference Price Models", *Journal of Consumer Research*, 24 (Sept): 202-214.

7. Ronald W. Niedrich, Subhash Sharma and Douglas H. Wedell. 2001. "Reference Price and Price Perceptions: A Comparison of Alternative Models", *Journal of Consumer Research*, 28(Dec): 339-354.
8. J. Miguel Villas-Boas and Russell S. Winer. 1999. "Endogeneity in Brand Choice Models", *Management Science*, 45(10): 1324-1338. (Account for price endogeneity, but no unobserved consumer heterogeneity).
9. Chu, Junhong, Pradeep Chintagunta and Javier Cebollada. 2008. "[A Comparison of Within-household Price Sensitivity across Online and Offline Channels](#)," *Marketing Science* (account for both price endogeneity and unobserved consumer heterogeneity), 27(2): 283-299.
10. Chintagunta, Pradeep and Junhong Chu. 2010. "Category Selection and Store Choice Modeling", working paper.
11. Andrews, Rick L., Andrew Ainslie and Imran S. Currim. 2002. "An empirical comparison of logit choice models with discrete vs. continuous heterogeneity", *Journal of Marketing Research*, 39, 479-487.
12. Gönül, Füsun and Kannan Srinivasan (1993), "Modeling Multiple Sources of Heterogeneity in Multinomial Logit Models: Methodological and Empirical Issues", *Marketing Science*, 12 (Summer), 213-229.
13. Chintagunta, P.K., D.C. Jain and N.J. Vilcassim (1991). Investigating Heterogeneity in Brand Preferences in Logit Models for Panel Data, *Journal of Marketing Research*, 28, 4, 417-428

#### **Models on Partial Ranking Data and Exploding Logit on Ranked Data**

1. Beggs, S., Cardell, S. and Hausman, J. 1981. "Assessing the Potential Demand for Electric Cars", *Journal of Econometrics*, 16, 1-19.
2. Hausman, J.A., and Rudd, P. A. 1987. "Specifying and testing econometric models for rank-ordered data", *Journal of Econometrics*, 34, 83-102.
3. Ophem, H.V., P. Stam, B. V. Praag. 1999. Multichoice Logit: Modeling incomplete preference rankings of classical concerts, *Journal of Business Economics and Statistics*. 17 117-128.
4. Kapman, Randall G. and Richard Staelin. 1982. "Exploiting rank ordered choice set data within the stochastic utility model", *Journal of Marketing Research*, 16(August) 288-301.
5. Yang, Sha, Greg M. Allenby and Geraldine Fennell. 2002. Modeling Variation in Brand Preference: the Roles of Objective Environment and Motivating Conditions, *Marketing Science*, 21(1): 14-31.

#### **NBD model:**

Puneet Manchanda, Peter E. Rossi, and Pradeep K. Chintagunta. 2004. "Response Modeling with Nonrandom Marketing-Mix Variables", *Journal of Marketing Research*, Vol. XLI (Nov), 467-478.

**Pareto/NBD model** – purchase event/duration model, the purchases conditional on being alive.

Schmittlein et al. 1987. "Counting your customers: who are they and what will they do next", *Management Science*. This article develops the probability of a customer still being alive based on her transaction history.

#### **The discrete analog of Pareto/NBD model – Beta-geometric/beta-Bernoulli (BG/BB)**

Fader, Peter S., Bruce G.S. Hardie and Jen Shang. 2010. "Customer-base Analysis in a discrete-time non-contractual setting", *Marketing Science*, 29(6), 1086-1108.

#### **Timing Models (Regression with Censored Data)**

1. Greene. 2000. 4<sup>th</sup> edition. *Econometric Analysis*, chapter 20, "Limited Dependent Variables and Duration Models".

2. Rupert Miller and Jerry Halpern. 1982. "Regression with Censored Data," *Biometrika*, 69, 3, 521-31. (The paper describes and compares four regression techniques with censored data that do not assume particular parametric families of survival distributions: Cox (1972), Miller (1976), Buckley & James (1979) and Koul, Susarla & Van Ryzin (1981)).
3. Kalbfleisch, J. D. and Prentice, R. L. 1980. *The Statistical Analysis of Failure Time Data*, New York: John Wiley & Sons.
4. Nicholas M. Kiefer. 1988. "Economic Duration Data and Hazard Functions," *Journal of Economic Literature*, 26 (June): 646-679.
5. Dipak C. Jain and Naufel J. Vilcassim. 1991. "Investigating household purchase timing decisions: a conditional hazard function approach", *Marketing Science*, 10(1): 1-23.
6. P.B. Seetharaman and Pradeep Chintagunta. 2003. "The Proportional Hazard Model for Purchase Timing", *Journal of Business and Economic Statistics*, 21(3): 368-382.
7. Peter Boatwright, Sharad Borle, Joseph Kadane. 2003. "A Model of the Joint Distribution of Purchase Quantity and Timing", *Journal of American Statistical Association* 98, 2003; 564-572.
8. Andre Bonfrer and Xavie Dreze. 2009. "Real-time Evaluation of E-mail Campaign Performance", *Marketing Science*, 28(2): 251-263.
9. Helsen and Schmittlein. 1993. Analyzing Duration Times in Marketing: Evidence for the Effectiveness of Hazard Rate Models, *Marketing Science*, 11(4), 395-414.

### **Joint Models on What, When and How Much Decisions**

#### **Household Data**

1. Gupta, Sunil. 1988. "Impact of Sales Promotions on When, What, and How Much to Buy", *Journal of Marketing Research*, 25 (Nov), 342-355.
2. Chiang, Jeongwen. 1991. "A Simultaneous Approach to the Whether, What and How Much to Buy Questions", *Marketing Science*, 10(4): 297-315.
3. Chintagunta, Pradeep. 1993. "Investigating Purchase Incidence, Brand Choice and Purchase Quantity Decisions of Households", *Marketing Science*, 12(2): 184-208.
4. Boatwright, Peter, Sharad Borle and Joseph B. Kadane. 2003. "A Model of the Joint Distribution of Purchase Quantity and Timing", *Journal of the American Statistical Association*, Vol. 98, No. 463, pp. 564-572.

#### **Discrete-continuous/Multiple-discreteness/Multiple Discrete-continuous models**

1. Dubin, J.A. and D.L. McFadden .1984. An econometric analysis of residential electric appliance holdings and consumption, *Econometrica*, 52(2), 345-362.
2. Hannemann, M. 1984. The discrete/continuous model of consumer demand. *Econometrica*, 52, 541-561.
3. Hendel, I. 1999. Estimating multiple-discrete choice models: An application to computerization returns. *Review of Economic Studies*, 66, 423-446.
4. Dube, J.P. 2004. Multiple discreteness and product differentiation: Demand for carbonated soft drinks, *Marketing Science*, 23(1), 66-81.
5. Chiang, J. 1991. A simultaneous approach to whether to buy, what to buy, and how much to buy, *Marketing Science*, 10(4), 297-314.
6. Chintagunta, P.K. 1993. Investigating purchase incidence, brand choice and purchase quantity decisions of households, *Marketing Science*, 12, 194-208.
7. Arora, N., G.M. Allenby, and J.L. Ginter. 1998. A hierarchical Bayes model of primary and secondary demand, *Marketing Science*, 17, 29-44.
8. Abdul Rawoof Pinjari, Chandra Bhat. 2010. A multiple discrete-continuous nested extreme value (MDCNEV) model: Formulation and application to non-worker activity time-use and timing behavior on weekdays, *Transportation Research Part B: Methodological*, 44(4),562-583.

9. Chandra R. Bhat. 2008. The multiple discrete-continuous extreme value (MDCEV) model: Role of utility function parameters, identification considerations, and model extensions, *Transportation Research Part B: Methodological*, 42(3), 274-303
10. Chandra R. Bhat, Marisol Castro, Mubassira Khan. 2013. A new estimation approach for the multiple discrete-continuous probit (MDCP) choice model, *Transportation Research Part B: Methodological*, 55, 1-22
11. Marisol Castro, Chandra R. Bhat, Ram M. Pendyala, Sergio R. Jara-Díaz. 2012. Accommodating multiple constraints in the multiple discrete-continuous extreme value (MDCEV) choice model, *Transportation Research Part B: Methodological*, 46(6), 729-743
12. Chandra R. Bhat, Sudeshna Sen. 2006. Household vehicle type holdings and usage: an application of the multiple discrete-continuous extreme value (MDCEV) model, *Transportation Research Part B: Methodological*, 40(1), 2006, 35-53
13. Chandra R. Bhat. 2005. A multiple discrete-continuous extreme value model: formulation and application to discretionary time-use decisions, *Transportation Research Part B: Methodological*, 39(8), 679-707

### **Aggregate Data**

#### **The AIDS model:**

Deaton, Angusdeaton and John Muellbauer. 1980. "An Almost Ideal Demand System", *American Economic Review*, 312-326.

Salvo, Alberto. 2009. "Cut-throat fringe competition in an emerging country marketing: tax evasion or the absence of marketing power", the journal of industrial economics, Vol. LVII, No. 4, 677-711. A two-stage model: top stage captures overall demand for soft drinks, specified as a log-linear model of Q on Price Index; bottom-stage is an AIDS demand function on soft drink expenditure from the top stage.

#### **Log-linear model:**

Hoch, Stephen J., Byung-Do Kim, Alan L. Montgomery, and Peter E. Rossi. 1995. "Determinants of store-level elasticities", *Journal of Marketing Research*, 17-29.

### **Aggregate Logit Models**

1. Chintagunta, Erdem, Rossi and Wedel. 2006. "Structural Modeling in Marketing: Review and Assessment", *Marketing Science*, Vol. 25, No.6: 604-616.
2. Berry, S., J. Levinsohn, and A. Pakes. 1995. "Automobile prices in market equilibrium", *Econometrica* 63: 841-890. (Note: though widely cited, this is a very hard-to-read paper. A beginner can start with Nevo (2000, 2001) as below).
3. Nevo, Aviv. 2001. "Measuring Market Power in the Ready-to-Eat Cereal Industry", *Econometrica*, Vol. 69, No. 2, pp. 307-342. (Note: an easier to read version of this paper is "A Practitioner's Guide to Estimation of Random Coefficients Logit Models of Demand," *Journal of Economics & Management Strategy*, 9(4), 513-548, 2000. You can download the Matlab code from Nevo's website: <http://www.faculty.econ.northwestern.edu/faculty/nevo/supplements/supplements.html>)
4. Sudhir, K. 2001. "Competitive pricing behavior in the auto market: A structural analysis", *Marketing Science*, 20(1): 42-60.
5. Petrin Amil and Austin Goolsbee. 2004. "[The Consumer Gains from Direct Broadcast Satellites and the Competition with Cable Television](#)", *Econometrica* 72(2), 351-381. (Note: This is an application of BLP to individual level data. The authors use probit model. Similar application to individual level data is Chu, Chintagunta and Cebollada. 2007. "A comparison of within-household price sensitivity across online and offline channels", *Marketing Science*, forthcoming, where the authors use logit models.)

6. Chintagunta, Dube and Singh. 2003. "Balancing profitability and customer welfare in a supermarket chain", *Quantitative Marketing and Economics*, 1(1): 111-147.
7. Junhong Chu, Pradeep Chintagunta and Naufel Vilcassim. 2007. "[Assessing the Economic Value of Distribution Channels – An Application to the PC Industry](#)", *Journal of Marketing Research*, Vol. 44, no.1, pp29-41, 2007
8. Chu, Junhong and Pradeep Chintagunta. 2009. "[Quantifying the Economic Value of Warranties in the U.S. Server Market](#)", *Marketing Science*, Vol 28, No.1.
9. Petrin, Amil and Kenneth Train. 2010. "A Control Function Approach to Endogeneity in Consumer Choice Models", *Journal of Marketing Research*, 47(1).
10. Chu, Junhong. 2013. "Quantifying nation equity with sales data: A structural approach", *International Journal of Research in Marketing*.
11. Rossi, Peter. 2014. "Even the Rich can make themselves poor: A critical examination of IV methods in marketing applications", *Marketing Science*, 33(5), 655-672.

### Aggregated Nested Logit

1. Berry. 1994. "Estimating Discrete-choice Models of Product Differentiation", *Rand Journal of Economics*, 25(2): 242-262.
2. Cardell. 1997. "Variance Components Structures for the Extreme-Value and Logistic Distributions with Application to Models of Heterogeneity", *Econometric Theory*, 13, 185-213.
3. Hui, Kailung. 2004. "Product Variety under Brand Influence: An Empirical Investigation of Personal Computer Market", *Management Science*, 50(5): 686-700.
4. Breakers, Randy and Frank Verboven. 2006. "Liberalizing A Distribution System: The European Car Market", *Journal of the European Economic Association*, 4(1): 216-251

## V. Course Policy

**The University Regulations on academic dishonesty will be strictly enforced!** Academic dishonesty is behaviour in which a deliberately fraudulent misrepresentation is employed in an attempt to gain undeserved intellectual credit, either for oneself or for another. It includes, but is not necessarily limited to, the following types of cases:

- a. Plagiarism - The representation of someone else's ideas as if they are their own. Where the arguments, data, designs, etc., of someone else are being used in a paper, report, oral presentation, or similar academic project, this fact must be made explicitly clear by citing the appropriate references. The references must fully indicate the extent to which any parts of the project are not one's own work. Paraphrasing of someone else's ideas is still using someone else's ideas, and must be acknowledged. Please check the University Statement on plagiarism on the web: <http://www.hku.hk/plagiarism/>
- b. Unauthorized Collaboration on Out-of-Class Projects - The representation of work as solely one's own when in fact it is the result of a joint effort.
- c. Cheating on In-Class Exams - The covert gathering of information from other students, the use of unauthorized notes, unauthorized aids, etc.
- d. Unauthorized Advance Access to an Exam - The representation of materials prepared at leisure, as a result of unauthorized advance access (however obtained), as if it were prepared under the rigors of the exam setting. This misrepresentation is dishonest in itself even if there are not compounding factors, such as unauthorized uses of books or notes.

You are expected to do your own work whenever you are supposed to. Incident(s) of academic dishonesty will NOT be tolerated. Cheating or plagiarism of any kind would result in an automatic FAIL grade for the course plus strict enforcement of all Faculty and/or University regulations regarding such behaviour.



