



**Michigan
Technological
University**

Michigan Technological University
Digital Commons @ Michigan Tech

Dissertations, Master's Theses and Master's Reports

2016

A latent-class discrete-choice model to demonstrate how course attributes and student characteristics influence demand for economics electives: the challenge to increase enrollment

Reginald T. Dillingham II

Michigan Technological University, reggiedillingham@gmail.com

Copyright 2016 Reginald T. Dillingham II

Recommended Citation

Dillingham, Reginald T. II, "A latent-class discrete-choice model to demonstrate how course attributes and student characteristics influence demand for economics electives: the challenge to increase enrollment", Open Access Master's Thesis, Michigan Technological University, 2016.
<https://doi.org/10.37099/mtu.dc.etr/262>

Follow this and additional works at: <https://digitalcommons.mtu.edu/etr>



Part of the [Behavioral Economics Commons](#), and the [Education Economics Commons](#)

A LATENT-CLASS DISCRETE-CHOICE MODEL TO DEMONSTRATE HOW
COURSE ATTRIBUTES AND STUDENT CHARACTERISTICS INFLUENCE
DEMAND FOR ECONOMICS ELECTIVES: THE CHALLENGE TO INCREASE
ENROLLMENT

By

Reginald T. Dillingham

A THESIS

Submitted in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

In Applied Natural Resource Economics

MICHIGAN TECHNOLOGICAL UNIVERSITY

2016

© 2016 Reginald T. Dillingham

This thesis has been approved in partial fulfillment of the requirements for the Degree of
MASTER OF SCIENCE in Applied Natural Resource Economics.

School of Business and Economics

Thesis Advisor:	<i>Dr. William Breffle</i>
Committee Member:	<i>Dr. Matthew Kelly</i>
Committee Member:	<i>Dr. Latika Gupta</i>
School Dean:	<i>Dr. Dean L. Johnson</i>

“To my mother.”

Table of Contents

List of Tables	vii
List of Figures.....	viii
List of Equations	viii
Acknowledgements	ix
List of Abbreviations	x
Abstract.....	xi
Chapter 1: Introduction	1
1.1 The Development of the Study.....	3
Chapter 2: Literature Review: The Evolution of Discrete-Choice Models	6
2.1 Exploring the Problem.....	6
2.2 Student Preferences	7
2.3 The Development of LCA	8
2.3.1 <i>Discrete Choice Experiments</i>	8
2.3.2 <i>Random Utility Theory</i>	11
2.4 Preceding Latent Class Analysis	13
2.5 Latent-Class Analysis	14
Chapter 3: Survey Methods & Design	17
3.1 The Data & Survey Instrument.....	17
3.1.1 <i>Attributes & Covariates</i>	19
3.2 Designing a Discrete Choice Experiment.....	21
3.2.1 <i>Design Efficiency Criteria</i>	22
3.3 LGC Software.....	24
Chapter 4: Model specification, Estimation, and Results	25
4.1 The Latent Class Model.....	25
4.1.1 <i>Random Utility Model</i>	26
4.2 Estimating Parameters	29
4.2.1 <i>Maximum-Likelihood and the E-M Algorithm</i>	29
4.2.2 <i>Marginal Utilities and Marginal Rates of Substitution</i>	30

4.2.3	<i>Information Criterion Tables & Class Estimation</i>	30
4.3	Results	33
4.3.1	<i>Selecting the Final Model</i>	33
4.3.2	<i>Interpreting the Final Model</i>	35
	<i>Model Explanatory Power</i>	36
4.3.3	<i>Model Estimation</i>	38
4.3.4	<i>Student Tradeoffs</i>	41
4.4	Class Descriptions	42
4.5	Student Awareness Analysis	45
Chapter 5:	Discussion	47
5.1	Attracting Class-specific Individuals.....	47
5.2	Study Recommendations	49
5.3	Increasing Student Awareness.....	51
	<i>Additional Recommendations</i>	52
	<i>Future Studies</i>	53
References	54
	Works Consulted.....	62
Appendices	66
	Appendix A: MTU Enrollment Numbers	66
	Appendix B: Investigating Random Choice Responses.....	67
	Appendix C: Survey Versions 1-3	68
	Appendix D: Latent Gold Choice	75
	<i>Choice Format</i>	75
	<i>Conditional Probabilities</i>	75
	Appendix E: LGC Files: Design Matrix and Choice Sets.....	77
	<i>Choice Sets</i>	79
	<i>Alternative File</i>	79
	Appendix F: Student Awareness Correlations	82
	Appendix G: The Bootstrap Likelihood Ratio Test (BLRT)	84
	Appendix H: Parameter Estimates	85
	Appendix I: Descriptive Statistics.....	88
	Appendix J: Response Frequencies.....	90

<i>Summary of Survey Frequencies</i>	90
<i>NRE Awareness Questions</i>	90
<i>Attribute Response Frequencies</i>	92
Appendix K: Crosstabs Analysis	94
Appendix L: Sampled Courses Summary	95

List of Tables

Table 3.1: Discrete-Choice Questions	18
Table 3.2: Attribute Descriptions.....	19
Table 3.3: Covariate Descriptions.....	20
Table 4.1: Information Criteria LL functions	33
Table 4.2: 1-5 Class Models with No Covariates	34
Table 4.3: 1-5 Class Models with Covariates	34
Table 4.4: Attribute Preferences by Class.....	37
Table 4.5: Model Hit Rate	37
Table 4.6: Attribute Parameters	40
Table 4.7: Tradeoff between Class Focus and Class Time	41
Table A1: Enrollment in Natural Resource Courses (2009-2016).....	66
Table A2: Unconditional Probabilities	76
Table A3: Original Design Matrix.....	78
Table A4: Choice Sets	79
Table A5: LGC Alternatives Matrix	81
Table A6: Correlation Matrix of Questions 1-3.....	83
Table A7: Likelihood of Enrollment and Economic Interest Correlations	83
Table A8: Final Model Parameter Estimates	86
Table A9: Covariate Estimates	87
Table A10: Descriptive Statistics.....	88
Table A11: LGC Model Description Statistics	89
Table A12: Frequency Statistics	90
Table A13: Awareness Question 1 Frequencies	90
Table A14: Awareness Question 2 Frequencies	91
Table A15: Likelihood of Enrollment Frequencies	91
Table A16: Attribute Response Frequencies	92
Table A17: Awareness of NRE Master's Program & Students Citizenship Status	94
Table A18: Likelihood of Enrollment in Q1 or Q2 & Student Economic Interest.....	94
Table A19: Sampled Courses Summary	95

List of Figures

Figure 2.1: MTU NRE Enrollment	6
--------------------------------------	---

List of Equations

Eq. (1) LCM Probability Density Function	25
Eq. (2) Random Utility Model	26
Eq. (3) Discret Choice Random Utility Model	26
Eq. (4) Conditional Logit Model	27
Eq. (5) Likelihood Function.....	28
Eq. (6) Joint-Likelihood Function.....	28
Eq. (7) Latent-Class Likelihood Function	29
Eq. (8) Marginal Rate of Substitution.....	30
Eq. (9) Log-Likelihood Function Specification for Information Criteria.....	32
Eq. (10) Penalized Log-Likelihood Function	32
Eq. (11) Estimated Utility Function.....	39

Acknowledgements

It would have been nearly impossible to complete a thesis such as this without the help of several individuals. I would first like to thank my mother, Lori Andrade, and my father, Reginald T. Dillingham, for their continuous encouragement and support to further my studies into graduate school.

Next, I would like to give special thanks to my advisor, Dr. William Breffle, for the unprecedented amounts of personal and family time he set aside to provide knowledge and guidance from start to finish. Getting to know him on a more personal level the last few semesters has been a profound experience and pleasure. I consider my adopted fixation on the natural environment and economic theory in general to be a product of his demanding yet uniquely captivating teaching styles. Without his dispersal of knowledge on the thesis topic at hand and encouragement to scrutinize every important detail, this study would not exist. Thanks for everything Dr. Breffle.

Furthermore, I would like to thank my thesis committee members: Dr. Latika Gupta, for taking time out of her busy schedule to critique and edit my work; and Dr. Matthew Kelly, who took on my thesis upon his arrival at MTU, and whose insight provided an invaluable jump start in the data analysis process.

I would also like to thank Dr. Gary Campbell, my graduate coursework advisor, for always having an open door to answer any simple or complex questions I may have had throughout my graduate career.

I would like to thank Jennifer Thacher from The University of New Mexico for her involvement in the design matrix. I also thank Mr. Patrick Hufnagel for setting aside his own graduate studies to assist in Excel data input.

Finally, I would like to thank several professors in the economics department of MTU's School of Business & Economics for their easy-going participation in administering the survey, which underlays this entire study, in their respective courses. Without the cooperation of Dr. William Breffle, Dr. Gary Campbell, Dr. Latika Gupta, Mr. Bryan Lagalo, Dr. Paul Nelson, and Dr. Emanuel Oliveira, this study would not have produced such a large and quality sample.

Thank you all.

List of Abbreviations

Sample Adjusted BIC: (SABIC)
Best-Worst Discrete Choice Experiments: (BWCE)
Bootstrap Likelihood Ratio Test: (BLRT)
Choice Experiments: (CE)
Conjoint Analysis: (CA)
Contingent Valuation Methods: (CVM)
Discrete-choice Model: (DCM)
Independent and Identically Distributed: (IID)
Information Criteria: (IC)
The Institutional Review Board: (IRB)
Latent Class Analysis: (LCA)
Latent Gold Choice: (LGC)
Latent Class Model: (LCM)
Likelihood Ratio Testing: (LRT)
Marginal Utility: (MU)
Marginal Rate of Substitution: (MRS)
Maximum Likelihood: (ML)
Michigan Technological University: (MTU)
Multinomial Logit Model: (MNL)
Natural Resource Economics: (NRE)
Ordinary Least Squares: (OLS)
Probability Density Function: (PDF)
Random Utility Model: (RUM)
Random Utility Theory: (RUT)
Revealed Preference: (RP)
Stated Preference: (SP)
Willingness-To-Pay: (WTP)

Abstract

A primary goal of universities is to maximize student enrollment by improving course curriculum and enhancing specific programs. This is especially a challenge for smaller universities who want to offer specialized and highly diverse electives. This study aims to increase the quality and relevance of electives offered by understanding specifically what attributes students prefer more or less when choosing among alternatives. The results are used to explore how to use limited marketing and student-outreach financial resources to target students that are most likely to enter and complete courses and programs, based in part on their socioeconomic or demographic characteristics. The application is aimed at the economics unit at Michigan Technological University, which offers two programs: an undergraduate economics major and a master's program in applied natural resource economics. Using an efficiently-designed discrete-choice experiment, this study elicits over 700 students' stated preferences over a variety of attributes of economics courses related to the natural environment. Students were surveyed, and each student was presented with six different pairwise choice options that were developed based on 36 different alternative courses consisting of 8 attributes, such as class topics, professor rank, time of day, and research requirement. There were three significant covariates (political preference, economic interest, and domestic/foreign status). A latent-class discrete-choice random-utility model is estimated to probabilistically group students into different preference classes. Four preference classes emerge from the results that are highly heterogeneous in terms of the marginal utilities and the probabilities of being in a given class. For example, the largest class (40%) is made up of mostly conservative students, and the smallest class (12%) is made almost entirely of liberal students. While this study and the applicability of the specific results is unique to Michigan Technological University, the use of stated preference surveys and latent-class models is highly flexible and can be applied to any program at any university.

Chapter 1: Introduction

In the United States alone, over two thousand colleges and universities compete for increasingly scarce revenues to conduct academic programs. These academic institutions range from two-year technical colleges and small liberal arts colleges and universities that only offer undergraduate education; to giant universities that cater to a variety of undergraduate and graduate students up to the doctoral level. All of these institutions conduct outreach programs and configure course curriculum typically in an informal fashion, in hopes that will enhance enrollment and populate classrooms. This can be a challenging task, especially for smaller institutions that wish to offer highly specialized courses and majors in unique or nontraditional fields. This study applies a more ad hoc approach to particular course design by understanding the demand of students. One example of these programs, studied in detail in this experiment, is the applied natural resource economics (NRE) program, offered by Michigan Technological University (Michigan Tech).

A wide variety of environmental issues are affecting the social welfare of modern society. Some of these issues include mining industries, climate change, natural resource stocks, water purity, and ecologically-environmentally driven policies, to name a few. These areas merit the social benefits of a study that focuses on the natural environment. Furthermore, in order to positively affect change and influence equity regarding these environmental topics, college courses on economics of the natural environment can attract the innovative minds of students who will ultimately design solutions. To appeal to these students and influence them to further their studies and research in these areas, it is critical to understand individual student preferences and the underlying course attributes that attract students and those that deter them. While it may be easier for larger universities, like the University of Wisconsin-Madison for example, which has extensive resources to reach most varieties of student preferences, smaller universities can greatly benefit from marketing and curriculum plans to target prospective students and reasonably incorporate their underlying course preferences.

The purpose of this research is to use survey techniques and econometric models to examine and understand the preferences of a highly-heterogeneous student body. As a result, outreach programs can be conducted and course attributes can be configured that attract students and improve enrollment. While such studies using similar econometric models are copious in the areas of transportation, marketing, and environmental and health economics, no study has ever been conducted to estimate the preferences of and maximize the utility of college students with respect to their demand for course offerings. As such, this study is highly experimental but offers considerable potential benefits not only students but also to professors and their institutions to not only protect but grow their small, specialized programs and provide greater variety. This study is based on the preferences of students enrolled in economics courses at Michigan Technological University. As such, the study demonstrates how current students trade off various course attributes when they pick their courses once they are at the university. Investigating these course attributes can provide sufficient evidence of some direct areas of courses and programs, that if improved, can increase the total number of students enrolled. Albeit, the population represented is current students in economics classes only. It is certainly conceivable to consider larger populations, but that is beyond the scope of the current study. Consequently, while the study cannot predict the probability that students will take economics courses on the natural environment, it definitely can estimate how that probability would change in percentage terms if courses were changed: up or down; and a little, a lot, or not at all.

This study is grounded in utility-maximization microeconomic theory under the assumption that currently-enrolled students will choose courses that maximize their satisfaction, given the constraint that it is impossible for any student to enroll in his or her “perfect” course schedule. Resources in higher learning, like in any other business, are constrained and dictate what courses can or cannot be offered during any given semester. To do this, virtually all students enrolled in all economics classes at Michigan Tech were surveyed about their preferences and their characteristics on the first day of class of spring semester 2016 using a pencil and paper survey. This yielded a massive sample in excess of 700 students at zero cost because there was a captive audience.

The survey has several parts and purposes, but the most important aspect of the survey is that it uses a “stated-preference” (SP) method to gather information. SP surveys simply ask the respondent to say something that explicitly or implicitly shows us what they like and dislike. “Revealed-preference” approaches are the alternative and instead examine actual behavior rather than hypothetical choices. Because of the many thousands of configurations of student schedules along with the burdensome and costly data collection required to obtain those data, an SP approach was deemed the most plausible.

1.1 The Development of the Study

The survey first collected data on the relationship between student familiarity with economics field electives and their likelihood of taking courses or obtaining a graduate degree, thereby establishing the correlation between awareness and involvement. The survey also collected data on socio-economic and demographic variables, to determine whether such factors affect preference. But the most important part of the survey is a set of six choice sets, where each set is different and contains two alternative hypothetical economic courses possibly differing in terms of up to eight course attributes. The student is asked simply to choose which course he or she would choose if in a situation where one of the courses must be picked, regardless of whether the student likes or dislikes one or both courses. The universe of choices contains 36 alternatives and thus 18 choice pairs across three survey versions. Asking these “choice questions” (which is sometimes called “conjoint analysis”) repeated times across hundreds of students, where each pair is different, generates thousands of data points to use in an econometric model to estimate a utility function. Marginal utilities (MUs) and marginal rates of substitution (MRS) for key class attributes (e.g., focus of the class, research component, time of day) are estimated, and can be used to determine how the likelihood of enrolling in any given course will change in percentage terms as the course attributes change. For example, one might be able to fill morning classrooms, a time of day that students generally avoid, by offering a class with other attributes that may be appealing, such as a moderate level of effort.

To estimate the utility parameters to make such recommendations possible, the survey data must be used in conjunction with an econometric model that will generate quantitative results. Of course, no model can ever explain all of the choice of all of the

subjects, thus a discrete-choice random utility model (RUM) is used because students make discrete choices over a binary choice set. The literature on this type of model and the survey needed to collect the data for such a model are discussed in great detail in Chapters 2 and 3. As stated, this study application is unique, so the literature referenced largely has an oblique relationship with the task at hand. The technical details of the model, its estimation, and key results are presented in Chapter 4. Finally, a lengthy set of recommendations to adopt course changes to increase enrollment based on the model parameters concludes this study in Chapter 5.

Before moving on, it is worthwhile to discuss a few elements of the model. It has already been established that SP data will be used in a RUM. A logit formulation is specified because of the ease of estimation. One aspect of the sample of students taking economics classes that became apparent at the outset of pretesting is that student preferences significantly vary. Because of this, a model of heterogeneous preferences was specified; in which, different students are allowed to have different preferences. There are numerous ways to allow for varying preferences, several of which are discussed for other studies in Chapter 2. The most modern, state-of-the-art model to estimate heterogeneous preferences relies on the notion that most preference differences cannot be observed by the analyst as a function of basic socio-economic data, although these variables may provide some assistance in estimating heterogeneous preferences. Rather, economists prefer “random parameter” models, where the preference parameters vary randomly, but the distributional parameters of the probability density function (PDF) underlying the heterogeneous preference parameters are unobservable but can be estimated. Random parameter models have existed for decades, and started with parameters that were continuously distributed (e.g. normal or lognormal distribution). A host of issues arose primarily associated with the structure imposed by this assumption. Continuously distributed random parameters have been largely replaced for at least the last decade by a model called a Latent Class Model (LCM). This model specification still retains the desirable aspect that researchers can only estimate parameters probabilistically and not deterministically, but it adopts a discrete PDF, where respondents are put into clusters or “classes” instead of having continuously distributed preferences. This aligns well with a

multitude of evidence that consumers tend to congregate into groups. The LCM estimates not only the parameters for multiple classes (There are usually a small number of classes.), but is also estimates the class size and the optimal number of classes. Class membership may also be a function of other observable variables (socio-economic variables are the most popular). Finally, given the estimated characteristics of preference classes, different course configurations and outreach efforts can be targeted at each of the classes.

While the results and details of this particular application are interesting, specifically at Michigan Tech, they are not the main contribution of the paper. The method employed here could be used at any college or university and in any discipline or field. It is not relegated to just economics, and the application to courses in the natural environment is only an illustration of how it is done.

Chapter 2: Literature Review: The Evolution of Discrete-Choice Models

2.1 Exploring the Problem

The primary concern addressed in this study, focusing on curriculum planning for economics courses, arises due to undergraduate and graduate enrollment numbers in specific areas. The data in Appendix A are plotted as a time series in Figure 2.1.

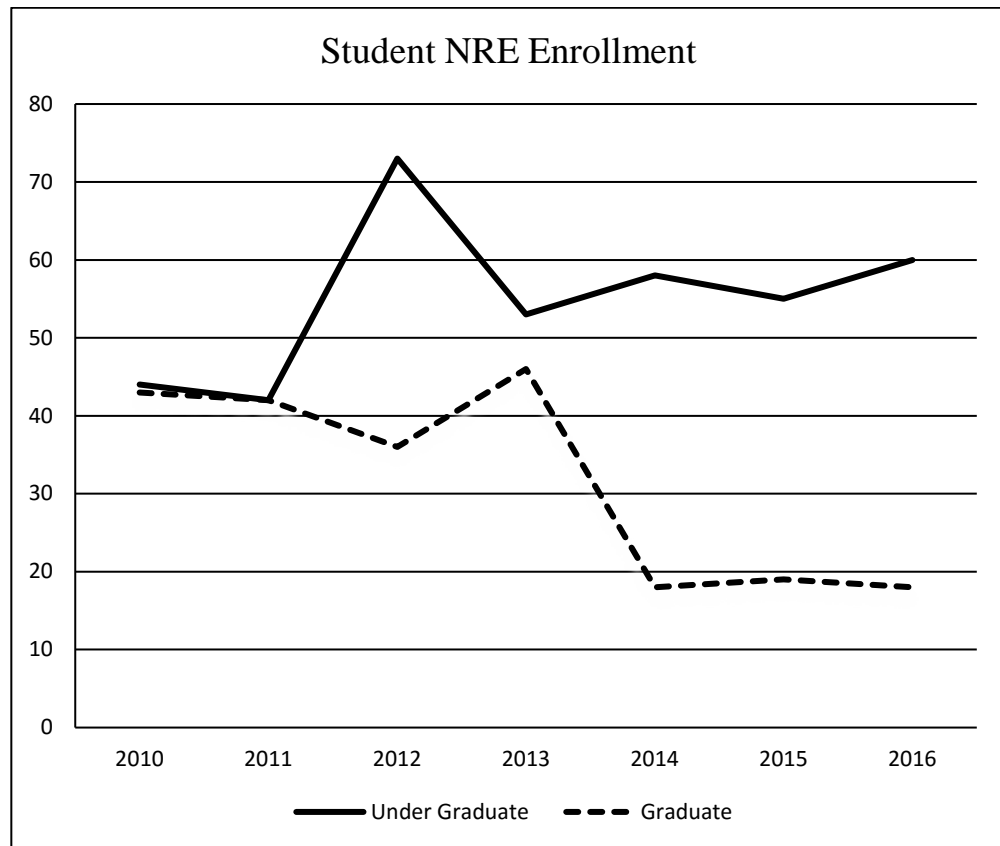


Figure 2.1: MTU NRE Enrollment

Since economic research has never used discrete-choice latent-class experiments to reveal heterogeneity in student preferences, the next section draws on what researchers have done, and makes inferences broadly designed to examine student preferences and identify course designs. The most similar research to this particular methodology in

academia is found in a working paper by Morey and Thacher (2012). They conduct a study on academic economists' preferences in which they use discrete-choice latent-class choice modeling to determine how respondents would choose among different hypothetical faculty positions as a function of teaching load, department rank, salary incentives or reductions, as well as a few other attributes. They conclude a four class model fit the data best and that the two large classes are similar while the smaller two classes varied significantly.

2.2 Student Preferences

Students typically enroll into a college or university before making specific choices about their course scheduling. Furthermore, some students do not even decide on a major until they are upperclassmen. For the current study, analyzing revealed-preference behavioral data through revealed preferences is impractical in evaluating student enrollment choices because interpretable data on full course schedules are not readily available or easily analyzed. However, the field of non-market valuation provides some viable methods that can be used to not only identify student preferences, but also classify students into finite consumer preference groups. The SP data in this context is generated by hypothetical tradeoff scenarios to determine what students want and how their preferences vary. There has been some work done in the area of exploring student preferences.

For example, Yang and Tsai's (2008) study of student's preferences on web-based learning finds that individuals prefer classroom-like simulations of online courses: Maringe (2006) conducted a study of UK students' preferences in deciding their university and courses of study using Likert-scale attitudinal questioning.¹ Results show students take a "consumeristic approach" when choosing whether or not to further their studies (i.e. benefits can be measured in terms of money). Further, the study shows evidence that students prefer courses of study that will best position them in the labor market rather than

¹ A scale based questionnaire method where respondents answer based on their level of agreement or disagreement to the underlying question.

choosing to enroll in post-graduate studies, disregarding whether they are interested in or have a passion for their major. Major improvements to the underlying graduate programs are suggested.

Other previous studies show that there are differences in preferences for student seating arrangements based on course types (McCorskey and McVetta, 1978). For a required course students preferred a traditional row seating arrangement; and for elective courses it was found that majority of students prefer an alternative horseshoe seating arrangement. Students also seem to have a preference towards courses taught by PowerPoint presentations (see Apperson et al., 2006). The respondents perceived PowerPoint presentation taught courses as more organized and efficient, correlating to much higher utility for courses taught with the use of PowerPoint as a key part of lectures.

These examples are used to illustrate some work done on analyzing distinct preferences of students. Most of these studies generated student preferences through alternative methodologies, thus the subsequent sections in this chapter focus on the development of the procedures relative to the underlying experiment, including random utility theory (RUT), discrete-choice models (DCMs), and latent class analysis (LCA).

2.3 The Development of LCA

2.3.1 Discrete Choice Experiments

Discrete-choice experiments, sometimes referred to as just “choice experiments” (CEs), are a type of stated preference method. CEs can be applicable in various situations, especially where consumption (typically goods) is some function of individual or consumer preferences. While the use of CEs is not new, the application of a CE in the context of student preferences is. Though different design variations exist (e.g. dichotomous choice, inclusion of an opt-out choice, or attribute rankings) their applications vary based on the researcher and type of study being conducted. For example, one reason for the use of dichotomous choices in this study is that time and were very limited (see Chapter 3).

Economists in the last decade have begun to predominately shift from the previously dominant contingent valuation methods (CVM) to CEs for environmental

evaluation of environmental disincentives (Breffle, 2008).² CEs have predominately been used for marketing purposes on the basis of consumer-welfare theory. Other applications include environmental, health, and transportation economics (Carlsson et al., 2003; Ryan and Gerard, 2003; Ben-Akiva and Boccara, 1995). CE's are preferred to CVM for a variety of reasons, one being that the former provides a much more flexible experiment to value "any" scenario once preference parameters are estimated. The latter is aimed at obtaining a direct statement of benefits from one, or at most a few, underlying projects.³ A CVM study could be used just to obtain a mean of willingness-to-pay (WTP) estimates for a certain project without any type of model estimation. CEs can map out parameters of an entire utility function to indirectly estimate utility. Additionally, it is harder for respondents to bias responses in CEs because they are asked indirectly about preferences and not directly about value.

CEs have long stood as a reliable data-collection process to use in conjunction with random utility theory (RUT) and economic demand theory (Louviere, 2006). As such, relating this study to common applications would mean considering this as a marketing application for economic courses on the natural world. Common applications of CEs include marketing for retailers to explore consumer preferences of different products. This is used to maximize utility of potential consumers and measure the marginal rates of substitution, alternatively known as the tradeoffs consumers make between different alternatives (Arcidiacono et al., 2012 and Lancsar and Louviere, 2006). Lockshin et al. (2006) utilize CEs to elicit the preferences of consumers when choosing different wines. A recent study in economic markets by Ortega et al. (2015) use CEs to assess the consumer food quality preference across a few different international retail channels and emphatically conclude a shift towards preferences for animal welfare.

CEs have also been used to evaluate diagnostic, treatment, and preventive medical errors in health care (Cunningham and Geller, 2011). Ryan and Gerard (2003) apply a CE

² See Hanley, Mourato, and Wright (2001) for literature exemplifying CVM.

³ CVM can be direct or indirect depending on payment vehicle used; but it is typically direct.

to healthcare benefits through labor characteristics to better understand respondents' preferences and interpretation of different attributes in healthcare economics (e.g. price). A recent CE by Hill et al. (2012) compares women's preferences for invasive versus noninvasive prenatal down syndrome tests. Their preferences were found to differ greatly. Lanscar and Louviere (2006) discuss the rationality of omitted responses in discrete choice experiments in health economics and suggest that such absent responses can still fit RUT. Lanscar et al. (2013) conduct a study on 'Best-Worst Discrete Choice Experiments' (BWCE), a more recent form of CEs. BWCE ask respondents to select their favorite (best) alternative and then the worst alternative to get richer insight on the attributes themselves.

Use of CEs in health economics have been increasing over the last decade, (e.g. Chestnut et al. (2012); Lanscar et al., 2013; Whitty et al., 2014; Louviere and Lanscar, 2006). See Clark et al. (2014) for a complete review of CEs in health economic literature from 2009-2012.

Further recent applications of CEs have been seen in labor supply (e.g. Gibson et al., 2015). Thoresen and Vatto (2015) estimate a discrete choice labor model and compare alternatives generated on tax reform to validate the labor supply models. Another recent study piloted by Behrman et al. (2013) modeled "teacher and non-teacher labor supply decisions to explore how wage policies affect the composition of the teacher labor force in both public and private schools" (Behrman et al., 2013). CEs are additionally useful in the transportation industry to uncover individual travel preferences or to discover alternative fuel preferences (Hackbarth and Madlener, 2013; Tanaka et al., 2014). Tanaka et al. (2014) conclude that both the U.S. and Japan are dominated by traditional conventional gasoline fueled vehicles. Daziano et al. (2013) provide a recent illustration of transportation networks and the issues in road safety for analyzing criteria such as accident risk. For more specific framework on constructing choice models to elicit transportation preferences see Ben-Akiva et al. (2013).

Other applications of CEs are seen in environmental economics (Adamowicz, 2004; Breffle and Maroney, 2009; Scarpa et al., 2007, 2006; Hoyos et al., 2015); where it is typically used to estimate the monetary values for ecological services, injuries to public amenities, and abatement projects (see Breffle and Maroney, 2009). Studies such as

Michaud et al. (2012) use CEs to elicit WTP for agricultural products whereas some other environmental choice experiments aim to discover the value of landscape improvements (e.g., Campbell et al., 2006). See Johnson et al. (2013) for a complete empirical guide for designing CEs in environmental economics.

As previously mentioned, this study is also rooted in consumer economics welfare theory that is modeled based on RUT, discussed in the subsequent section. Econometric models can be specified to estimate quantitatively how students are willing to tradeoff course attributes. This provides technical, not just heuristic, information about how changing attributes will affect the overall probability of enrollment using a maximum likelihood (ML) function and how that probability varies across students via “heterogeneity,” through a LCM. However, before doing so a random utility model (RUM) must be constructed to estimate individual utility parameters based off the discrete-choice responses.

2.3.2 Random Utility Theory

RUT assumes that there is an unobservable, random component that reflects the influence of all other excluded variables that affect choice. Utility can be defined as the individual welfare or satisfaction gained or loss from making a decision between two or more alternatives. The human decision making process generates choices that maximize utility; the choices that individuals’ make are based on a finite set of alternatives. Students, like consumers, will choose alternatives from a discrete number of choices that maximize their utility. The basis of this study is rooted in McFadden (1973), who suggests that there is a hidden (latent) property in all functioning humans called utility. This is an ordinal component that cannot be readily observed by researchers (McFadden, 1973; Louviere et al., 2010). It is assumed individuals make choices in order to maximize their own utility; and thus, their choices can be modeled to show us interpretable parameters of their utility.

That is, people, including college students, make choices with the purpose to maximize their own utility. Written as $U_j \geq U_{j_n}$ for all $j \in J_n$ where individual i chooses j if and only if the utility gained from U_j is greater or equal to the utility gained from alternative U_{j_n} . J_n is the set of alternatives faced by an individual i .

Since choice, the dependent variable, is not a continuous variable, linear regression is not appropriate. Instead, the development of a logit model based on individual behavior rules is constructed to estimate the probability of an individual making a choice.⁴

In this experiment a discrete-choice RUM is specified where study subjects (i.e., students) provide their choices over a set of alternatives; and the model is used to estimate parameters that best explain those choices. From these choices the parameters of a utility function can be estimated, and marginal utilities (MUs) for various preference groups can be calculated along with marginal rates of substitution (MRS). MRS are equal to the slope of indifference curves: students' willingness to tradeoff attributes hold utility constant.

RUMs have traditionally consisted of two components that are: (1) explainable and (2) unexplainable, put simply. These are also referred to as systematic or deterministic versus random components (Louviere et al., 2010). Systematic components are those choices that can be readily identified, assumed, and explained based on model parameters. On the other hand, random components (ϵ) are all the unexplainable influences of choice and consist of some unobservable utility associated with that individual choice. Later, the deterministic component is allowed to have random characteristics as well.

The ϵ term denoted above is a stochastic error term that reflects distinct idiosyncrasies of the specific individual (McFadden, 1973). This random term accounts for the variance in taste and preferences that are random in nature.

No model will predict all choices with 100 percent accuracy. Therefore, to specify a DCM an assumption must be made about the distribution of this stochastic error component, ϵ . Typically, there are two options considered among econometricians; logit and probit.

Generally speaking, a probit model assumes the ϵ is standard normally distributed; [$\epsilon \sim N(0, \sigma)$]. A "logit" specification assumes that ϵ is independent and identically

⁴ McFadden (1973) describes H as a set of behavior rules that describe the population, such as a set of demand functions that maximize some utility function. Where h is the maximization of a "specific" utility function and contains all possible attributes in an alternative set.

distributed (IID) type 1 extreme value distributed [$F(\epsilon) = \exp(-e^{-\epsilon})$] (see McFadden, 1973). This study has a discrete-choice logit model. In practice, the choice across these distributions has minimal effect on results, however there are reasons why logit is preferred to probit. The primary reason the logit model dominates is because of the ease of integration. McFadden (1973) proved that the estimation of the probability of choosing one alternative over another is not a function of ϵ , no matter how many alternatives are in the choice set. That is to say that with logit there is a closed form solution and ϵ is IID rather than independent of irrelevant alternatives (IIA).⁵ LCMs relax the IIA assumption by making the error component correlated across alternatives. Conversely, probit requires numerical integration which becomes increasingly complicated with more alternatives.

Many researches have explored the randomness in individual response patterns using psychological and mathematical studies in order to attempt to explain choice behavior. A further review of this phenomenon can be seen in Appendix B.

The RUM, however, is the first step in the LCM development. The next section builds on this theory and explores preceding methods leading to the development of LCA.

2.4 Preceding Latent Class Analysis

Prior to the last couple of decades, the V expression in the RUM was strictly deterministic. There are dozens of older studies (e.g., Brock and Durlauf, 2001; Birol et al., 2006) that incorporate heterogeneity through the use of interactions between utility parameters and observable characteristics. There are however two main issues that arise when using the latter methodology. First, these simple interactions may introduce bias in the results if the specification of V is incorrect; and second, observable covariates may not produce useful results if much of the heterogeneity is unobservable or latent. McFadden (1973) posits the first discrete-choice random-utility model, and to date is the most famous and widely applied methodology for modeling choice behavior. There have been hundreds of models based on this original concept where a RUM has been applied to consumer choices (e.g. Roy et al., 1996; Swait and Adamowicz, 2001; Boxall and Adamowicz, 2002;

⁵ IIA restricts preferences across individuals to vary.

Kim et al., 2002). There are however restrictions on this original theory of McFaddens' that have since been modified to generalize some of the original assumptions of the model.

Earlier RUT logit models assumed that all preferences of individuals with the same characteristics were homogeneous. While this approach gained popularity, analysts argued that the distributional assumptions assumed for preference parameters were too restrictive (Dayton and Macready, 1988). Another issue arose because this did not comport well with classic observations that preferences across consumers often congregate into a finite number of preference classes. Train (1998), in his recreational demand model, took a more flexible and agnostic approach to account for choice heterogeneity by incorporating a continuously distributed random preference parameter into V . By doing so both V and ϵ became random and β parameters adopted a subscript i (β_i) allowing them to continuously vary across the sample. The LCM is more flexible and can handle a larger range of behavioral and categorical variables (Green and Heshner, 2003). As a result, attention turned to the use of LCM to incorporate heterogeneity as early as Green Jr. (1951).

2.5 Latent-Class Analysis

LCMs are becoming more and more prevalent for researchers who wish to model individual choice heterogeneity based on alternative attributes and characteristics. LCA puts individuals into probabilistic groups based on the likelihood that they belong in that group, and is the most flexible heterogeneous method among alternatives. Take cluster analysis for example, another relatively popular grouping method, but now quite outdated. When clustering preferences, strong assumptions are adopted that puts people in preference groups with certainty, opposed to probabilistically. This could ultimately bias parameters. LCA does not put individuals in discrete bins but instead allows for the possibility that individuals could be in any bin. Each class has differing marginal utility levels for attributes. For example, LCA can predict liberals are most likely to be in one class but allows the possibility that a liberal could be in another class with a much lower likelihood. As Vermunt and Magidson (2005) explain, a typical violation in early RUMs comes from the assumption that the researcher's observations of individual i are independent. Thus, incorporating random parameters, as this study does, modifies this assumption and allows

for choices to be correlated, as they are in reality. Further analysis of the LCA literature is provided subsequently.

Latent class theory assumes that the population consists of C total preference classes where membership in a specific class is based on utility from the attributes and respondents' characteristics. An individual's true class is unobservable (i.e. latent) by researchers. Campbell et al. (2011) describe the theory of latent-class modeling as "individual choice behavior depending on observable attributes and latent heterogeneity that varies with different factors that are unobservable by the analyst." The probability that an individual will be in class c can also be estimated as a function of covariates, which in this study are sociodemographic characteristics of the student. The preference class probabilities must sum to 100%, and all vary in size.

LCMs essentially "tease" out the MU parameters that are associated with different classes. Within any given class, these preferences are homogenous, but conversely these preferences differ between classes. It is possible, through LCA, to empirically determine a set of influential forces that cause certain individuals to make choices the way they do. There are many applications of LCA that can be used for a variety of empirical purposes, from transportation data to experimenting and modeling consumers and the preferences that drive their purchases. For example, Briesch et al. (2012) apply a latent class model to study grocery store preferences and the tradeoffs between lower prices and increased driving distance to the store. Boxall et al. (2002), in their environmental application, forecast demand of recreational parks using a LCM to create segments where membership is dependent on preferences of wilderness park attributes. Morey et al. (2006) use attitudinal data to estimate a LCM on angler preferences for fishing site characteristics. They concluded that while all anglers would prefer less PCBs in the fishery, their sample of anglers breaks into three significantly distinct preference groups.

Morey and Thacher (2012) conducted an experiment using methods similar to this study, where a discrete-choice latent class model was used to explore how academic economists' tradeoff between their actual faculty position with other attributes such as teaching load and salary. The study concluded on a four class latent class model where

preferences and marginal utilities differed throughout classes and only one covariate, average *cites/year*, was found to be significant (see Morey and Thacher, 2012).

Extending the health applications of LCA, some medical doctors have found LCMs useful for the improvement of surgical procedures, medical treatments, and diagnostics testing. Specifically, Qu et al. (1996) applied a LCM with two classes, diseased and non-diseased, to test the accuracy and sensitivities of biological diagnostics test in patients whether a disease is or is not present (see Qu et al., 1996). This slightly differs from typical LCA health applications where the model is a function of patients or consumers' preferences because this study is not constructed based off of choice data. See Lanza and Rhoades (2013) for a more recent study of LCMs used in health economics.

As can be seen above, LCMs can be applicable in a variety of different scenarios where choice data are to be modeled. The next chapter expands on the design and implementation of the survey instrument used in this study, and also draws on some key CE design elements used to obtain the choice and covariate data needed in order to estimate a LCM.

Chapter 3: Survey Methods & Design

3.1 The Data & Survey Instrument

The data gathering process began with research of the primary attributes in economics courses offered at MTU. An email was then sent to all economics professors requesting them to participate in the study by administering this survey to all of their students in their specified courses. Pretest surveys were then designed and administered to a variety of individuals, both students and non-students, as well as all economics professors in the unit. Those pretests were then discussed with the individuals and ultimately led to modifications of the survey. Constrained by limited time due to professor lectures, the survey was designed to be completed in ten minutes or less. Thus, the survey instrument was limited to two sides of one piece of paper to be completed in the ten minutes.

The students were selected based upon their enrollment in all economics courses offered by the MTU School of Business & Economics in the spring semester. It is worth noting that because some students were enrolled in more than one economics course, they may have taken the survey more than once.⁶ It is uncertain because of the anonymous nature of the choice experiment. All professors agreed to participate, and surveys were hand delivered to each professor and returned in January of 2016. The survey uses pairwise choice options (see Table 3.1) partly due to faculty preference for a <10-minute survey to limit interference of class time, along with the large sample of >700 economics students taking the survey.

The survey instrument consists of two important question blocks in addition to the pairwise choices for a total of three sets of questions ascending as follows: (1) three Likert-scale start-up questions with the purposes of warming up the respondent to the survey and to investigate the relationship of familiarity and the likelihood of enrollment; (2) a sequence of six binary choice questions where respondents choose their most preferred (Table 3.1);

⁶ Sampling a student multiple times is not an issue. It means they are weighted more highly in the data, which is appropriate.

and (3) a set of eight questions to elicit socio-economic and demographic information about respondents. These data are used in the LCM as covariates to determine class probabilities. There are three versions of the survey instrument, differing only in terms of the choice questions. All three are presented in Appendix C.

Table 3.1: Discrete-Choice Questions

Q4 Please pick the alternative you prefer (course A or course B):		
<u>Attributes</u>		
Focus	Conservation/sustainability	Resource use/energy markets
Credit	Graduation only	Graduation only
Study hours	7 hours	3 hours
Teacher	Tenured professor	Instructor/lecturer
Paper	Small portion of grade	Small portion of grade
Time	Late afternoon	Late afternoon
Day	MWF	Tu/Th
Semester	Spring	Spring
I prefer option: (circle one)	A	B

The administration of the survey was anonymous to comply with university guidelines and the survey instrument was evaluated and approved by the IRB before field tested.⁷ Of the 750 surveys administered, 723 respondents returned them for a response rate of ~96%. Researchers should note that this high response rate is due to a pen and paper survey being implemented to students during a fixed block of time. Students of this setting are expected to generate large samples because the topic is of great interest to students. The sample size ($N = 723$) multiplied by the 6 choice questions ($N * 6$) generated > 4300 choices across alternative pairs, although a few choice questions were left blank. Section 3.3 and Appendix D provides more information on the software use and data entry process.

⁷ The Institutional Review Board (IRB) is a committee under the American Public University System (APUS). This committee is designed to oversee and approve the research of human subjects to ensure the proper legal and ethical guidelines are followed.

3.1.1 Attributes & Covariates

The two tables shown below contain course attributes and the socio-economic characteristic data collected. All of these variables, including covariates, were converted into 0,1 dummy (i.e., binary) variables. The choice sets consist of 8 attributes but break into 12 total binary variables because some variables have 3 levels. For example, in Table 3.2 the attribute *focus* has three levels: conservation/sustainability, balance use and conservation, and resource use/ energy markets. All three of those variables take on discrete values (0 or 1).

Table 3.2: Attribute Descriptions

Attribute Name	Attribute Description	Levels	Attribute Levels
<i>focus</i>	The focus of the course.	3	conservation/sustainability , balance use and conservation, resource use/energy markets
<i>major</i>	Whether or not a major requirement is fulfilled by enrolling in the course.	2	Major requirement fulfilled, graduation only
<i>studyhr</i>	How many hours will likely be required per week to studying.	3	3, 5, 7
<i>profrank</i>	The rank of the professor teaching the course.	3	instructor/lecturer, junior professor, tenured professor
<i>smpaper</i>	Grade percentage of the required term paper.	2	small, large
<i>time</i>	The time of day of the course.	3	Early Morning, Midday, Late Afternoon
<i>MWF</i>	The day of the week of the course.	2	MWF, Tu/Th
<i>fall</i>	The semester.	2	Spring ,Fall

Table 3.3: Covariate Descriptions

Covariate Name	Covariate Description	Levels	Covariate Levels
<i>STATUS</i>	Student's college year/rank.	3	Under-classmen, upper-classmen, graduate student
<i>MALE</i>	Sex of the respondent.	2	male, female
<i>USA</i>	Indicates whether the student is foreign or not.	2	domestic, foreign
<i>MAJOR</i>	The students current major focus.	3	HASS, STEM, other
<i>INTEREST</i>	Level of interest in the general field of economics.	3	low, medium, high
<i>PLANS</i>	Students post-undergraduate plans	3	grad school, job market, other
<i>JOBTYPE</i>	The type of job the student desires.	3	business, government, other
<i>POLPREF</i>	The students political preference.	2	liberal, conservative

The 8 covariates used in this study are presented in Table 3.3, however, only three were found to be significant in the final model.

Using logic and empirical reasoning, pre-testing, as well as past course experiences summarizes the development of all original variables. For example, political preference (*POLPREF*) was thought to be important because of the two main party's often contrasting beliefs on environmental policy. All variables went through an elimination process.

In order to further clarify some theory and methods governing the design of efficient CE's the next two sections will provide some explanation of why the design efficiency and other statistical design criteria are important to conduct a quality CE.

3.2 Designing a Discrete Choice Experiment

This section explains the general survey procedures of the underlying CE, and of choice models in general, as well as the importance of the “statistical design” in estimating random-utility choice models with efficiency.⁸ If the researcher’s goal is to construct a defensible CE it is essential to have alternative sets that optimize efficiency of the statistical design. One important note is that increasing sample size can offset a poor design; however, this is not the present case. This study has an efficient design along with a large sample size (Brefle, 2008).

Kuhfeld (2010) asserts that there are three efficiency measures that can quantify the matrix size: (1) A-efficiency (2) D-efficiency and (3) G-efficiency (see Kuhfeld, 2010; Brefle, 2008). In the present study D-efficiency is used. It is the most common of the three, and is based off of the geometric averages of the eigenvalues (Lusk and Norwood, 2005).⁹

Research and selection of attributes and alternatives is the first process in optimal design efficiency. This selection process will not be uniform in all CEs, and thus should vary with the underlying experiment using logic, economic theory, similar experiments, and proper judgment (Alpizar et al., 2003).

The efficiency criteria in statistical design focuses on 4 main principles (Huber and Zwerina, 1996; Brefle, 2008): (1) orthogonality, (2) level balance (see Kuhfeld, 2010 p.66 for a simple illustration), (3) minimal overlap, and (4) utility balance. These are explained more in the following section.

⁸ The statistical or experimental design can be considered as the arrangement of attributes and alternatives into column and vectors and row vectors, respectively. This is often displayed in matrix form.

⁹ Eigenvalues are scalars that can provide a size measurement of the design matrix. They are associated with the respective eigenvector in linear transformations. See Lusk and Norwood (2005) and Kuhfeld (2010) for further detail.

3.2.1 Design Efficiency Criteria

Orthogonality

This element increases design efficiency makes it easier for the analyst to determine effects of all variables by intentionally designing the attributes in the study to be uncorrelated (Brefle, 2008). As a result, this eliminates the risk of multicollinearity. The statistical design of the survey instrument includes 36 hypothetical courses that are completely “orthogonal.” Appendix E contains an orthogonal array for a main effects design.¹⁰ Orthogonal designs can be described as every pair of levels appearing with uniform frequencies across all pairs of factors (Kuhfeld, 2010). Orthogonal arrays on the other hand consist of orthogonality as well as level balance, explained next. Main effects designs, such as this one, are used for its simplicity over designs with interactions and its significant reduction in the total number of choice sets. This means that it is assumed that attributes do not interact with each other in utility space.

Level Balance

Level balance, in theory, is achieved when all the attribute levels appear with same frequency. For example, this would imply that all attributes have two levels (i.e. 0 and 1). The purpose of level balance is to maximize the statistical power of the experiment, and some researchers “falsely” use higher levels for more important attributes to increase estimation accuracy (Brefle, 2008). While some researchers follow the theory that all attributes must exhibit the same level in order to achieve level balance and efficiency maximization (Lusk and Norwood, 2005), Louviere et al. (2005); and Brefle (2008) conversely posit that attributes may take different levels as long as the number is some exponent of the same prime number. This study is in line with the latter research, in that our design consists of attributes with 2 and 3 levels. Additionally, the fact that the 3-level variables break into 2-level dummy variables in all cases further underscores that this study has no continuous variables; only discrete (see section 3.1.1).

¹⁰ A main effects design means that only the main attributes are uncorrelated (interactions may be correlated).

Minimal Overlap

Minimal overlap in general, means that combinations of attribute levels do not repeat in a single choice set (Huber and Zwerina, 1996). This in turns makes the choice sets more intricate and increases the cognitive burden on respondents. Minimal overlap is however violated when the inclusion of a “status quo” (i.e. no change) choice is incorporated into the questionnaire because it is repeated in every choice set. Another method (Bunch et al, 1996) proposes the shifting method may be an alternative to maximizing attribute comparisons but is also inconsistent with when inclusion of status quo is present.¹¹ This study however uses “forced choice,” meaning that the respondent must answer the question without status quo as a possible choice. Thus, minimal overlap is achieved in the statistical design. Furthermore, there is no such “status quo course” offered where attributes are homogenous for all course electives. Status quo is often used in referendum formats and is not relevant in this study because it typically incorporates a monetary element. Simply put, the referendum format is rejected in the interest of increasing efficiency (see McFadden, 1994).

Utility Balance

Utility balance increases the information available to accurately estimate MU levels by forcing respondents to make more difficult decisions (Bunch et al., 1996). Utility balance is achieved when the alternative choice sets are close in utility space for respondents (Brefle, 2008). Another concern is that an “opt-out” alternative gives the respondent an easy choice to revert to when being asked a difficult question (Champ et al., 2005). This ultimately minimizes the utilization of information gathered from a respondent’s choice selection. There is no opt-out alternative in this study’s discrete choice design to maximize question difficulty and achieve both minimal overlap and utility balance. Nonetheless, the respondent can always opt-out from a choice question by leaving it blank; however, there were very few of these in this study. The lack of blank choice

¹¹ Bunch et al. (1996) discovered this design strategy, which involves adding a constant to each attribute level to produce additional alternatives.

questions (0.003%) favorably suggests that the respondents understood the task at hand and could tradeoff the attributes provided in the survey.

3.3 LGC Software

Latent Gold Choice (LGC) is used for all estimations. Immediately following survey implementation, the surveys were retrieved from the respective professors and data input began. 723 observations were recorded in Microsoft Excel and SPSS in *IBM SPSS Statistics*. SPSS statistics was used for a great deal of statistical input and development of correlation matrices (Appendix F) and other statistical comparison tools such as frequencies and crosstabs (Appendices J&K, respectively). The analytical program used for the modeling is *Latent Gold Choice* (LGC) version 5.1.0. These data sets consist of the following: (1) choice data (each subject's responses to "all" survey questions); (2) an alternatives file (the design matrix of all the different alternatives); and (3) a choice set file (consists of 18 different choice sets corresponding to alternatives.) See Appendix E for files (2) and (3).

Chapter 4: Model specification, Estimation, and Results

4.1 The Latent Class Model

This model assumes that the probability of individual i choosing alternative j in choice set k is a function of the individual's class membership, which is denoted as c . In LC choice models probability density function associated with the response of individual i has the form:

Eq. (1)

$$P(y_i|z_i) = \sum_{c=1}^C P(c: z_i) \prod_{j=1}^J P(y_i: x_i|c)$$

where y_i denotes the full vector of choice responses maximizing utility of individual i . $y_i = 1$ if individual i chooses alternative j in choice set k , and 0 otherwise. Thus, $y_i = 0$ or 1 corresponding to the alternative chosen; which is represented as A or B in the survey. z_i are characteristics of individual i , and x_i is the matrix of attribute levels in individual i 's choice set.

$P(c: z_i)$ denotes the probability of being in class c , unconditional on y ; but varying with z_i . These unconditional class probabilities are the probability of being in class c as a function of individual covariates (e.g., political affiliation), and are incorporated in the model to form relationships between the covariates (z_i) and the subset of individuals in the underlying class.

The LCM assumes that parameters vary discretely and randomly across classes. As such, the β term (which contains the MU parameters) seen in Eq. (3), becomes β_c , which varies over C classes. C , the total number of latent classes in the model, can take on values $1 \leq c \leq C$. The LCM does not only estimate the β_c 's, but also the likelihood that any individual, i , belongs to any of the unobserved C classes. This will in turn determine the optimal number of utility maximizing preference classes as well as the overall size of each of the classes. The model elements that constitute the LCM are discussed further in this

chapter, building on the original RUM to ultimately specify the latent class likelihood function.

4.1.1 Random Utility Model

Assuming utility (U) consists of two components, the general RUM, in its simplest form is:

Eq. (2)

$$U_{ijk} = V_{ijk} + \epsilon_{ijk}$$

where U_{ijk} is the latent utility individual i , ($i = 1, \dots, N$) associates with choice alternative j ($j = 1, \dots, J_n$) in choice set k . U_{ijk} consists of two components: V_{ijk} , which is the non-stochastic, “deterministic,” element of utility. And ϵ_{ijk} which is the stochastic, “random,” latent component that reflects the uniqueness of individual i . V_{ijk} is the systematic utility that is a function of the explanatory variables as well as the unobserved random β parameters, which are the MUs. Assuming the latter, the RUM can be written as:

Eq. (3)

$$U_{ijk} = V_{ijk} + \epsilon_{ijk} = \beta' x_{ijk} + \epsilon_{ijk}$$

where β is a vector of unknown coefficients associated with students’ preferences that varies “randomly” over all students. x_i is a vector of attribute levels based on individual i ’s choices from the design matrix of all alternatives. Utility is “conditioned” on x_i , which are based on individual choices. Thus, Eq. (3) is a conditional logit model. x_{ijk} is a vector of explanatory variables for individual i and the j th alternative in choice set k . In Eq. (3) the β vector varies randomly over all students rather than remaining fixed as in the RUM presented in Eq. (2). As such, this is a RUM where preferences vary across individual i .

All parameters of $U(\cdot)$ are random. In this case, a preference parameter, β_{1i} for example, takes on a statistical distribution, and in a LCM that distribution is assumed to be discrete because it takes on a finite number of values.

The LCM generalizes the conditional logit model (McFadden, 1973) to incorporate heterogeneity through distinct preference classes and to generate parameter estimates that vary by class for changes in the attributes of the alternatives. In a given class, the correlation across an individual's choices are all attributed to the class membership. However, after conditioning upon class, all responses are considered independent across choice pairs. The conditional logit model used is essentially a regression model for choice data that determines the probability that individual i selects alternative j from choice set k

Eq. (4) is a probability model generalized to estimate the probability of individual i choosing alternative j in choice set k “conditional” upon in the individual's class membership. For an example specific to this study consider the following. In an A-B choice question, where $c = 2$, the conditional probability that A is chosen over B takes the binomial logit form.¹²

Eq. (4)

$$P(A) = P(U_{Aij|c} > U_{Bij|c}) = \left(\frac{e^{V_{Aij|c}}}{e^{V_{Aij|c}} + e^{V_{Bij|c}}} \right)$$

Likelihood Function

Given the form of the probability functions, a likelihood function, L , must be specified. The likelihood function (L) is simply a joint probability density function (PDF) made up of both choice probabilities conditional on y_i and unconditional class probabilities $P(c: z_i)$. The goal is to find the model parameters that maximize this function using maximum likelihood (ML) estimation. The different levels of covariates used helps to predict different class memberships using this likelihood function (Dayton and Macready, 1988).

Given the initial assumption of independence, Eq. (5) is an example of the likelihood of observing individual i 's response to a given choice question, “conditional” on their other choices.

¹² $|c$ is used to denote that probability is “conditional on membership of class c .”

Eq. (5)

$$L_i = \prod_i^N \prod_{j=1}^J \left(\frac{e^{V_{ijk}}}{\sum_{j=1}^J e^{V_{ijk}}} \right)^{r_{ijk}}$$

where r_{ijk} denotes an indicator that is equal to 1 if individual i picks alternative j in choice set k , and 0 otherwise.

Let P_{ijk} denote the probability that individual i will choose alternative j in choice set k shown in Eq. (6) below. A LCM such as this one assumes that $P_{ijk|c}$ is a function of c , where c is individual i 's group membership. Thus, the probability of the six choices occurring simultaneously is a product of their probabilities. This gives us the joint likelihood function for a latent class choice model:

Eq. (6)

$$L = \prod_i^N \left[\sum_{c=1}^C P(c: z_i) \prod_{j=1}^J P_{ijk|c}^{r_{ijk}} \right]$$

The likelihood function in Eq. (6) incorporates the unconditional probabilities of class membership, $P(c: z_i)$, and $P_{ijk|c}^{r_{ijk}}$ represents the probability of the individuals response patterns conditional on being in class c .

For each individual i and each class c , the likelihood function in Eq. (6) has two main components: the first is the probability of class membership given i 's covariates, and the second is the traditional likelihood function for the pairwise course choices conditional on being in class c . The first component can be thought of as a weight, where all of the weights across the C classes sum to one. The second is a joint probability density function of the utility parameters, the β 's, which explain the choice data by class. These parameters are also the MUs of the indirect utility function.

4.2 Estimating Parameters

4.2.1 Maximum-Likelihood and the E-M Algorithm

The likelihood function is maximized using the Expected Maximization (E-M) algorithm, which is a Bayesian process, along with a numerical search algorithm. As such, the model iterates many times until a global maximum is found, with continual updating until a solution is reached. The process starts with a vector of conditional class membership probabilities, conditional upon the choices and the covariates. These are averaged across the sample to obtain unconditional class membership probabilities that are no longer conditioned by the choices. This is the first probability expression in Eq. (6). Then the second maximum likelihood numerical search algorithm is used to find the values of the β 's that maximize the likelihood function. Then Bayes' Theorem is invoked to recalculate all of the conditional class membership probabilities, again conditional on both the covariates and the choices. That is the end of the first iteration. Iterations continue until the model parameters stop changing significantly and there are no further improvements to the likelihood function. A detailed discussion of all of the equations and their relationships to one another can be found in Breffle et al. (2011). This process is done by LGC. The results are easily interpretable parameters due to its ease of use and advanced features that supersedes other LCA software.

Because the product of probabilities is typically an extremely small number, maximizing the likelihood function can be difficult for the maximum likelihood search algorithm: small changes in the parameter estimates can cause large swings in the calculated likelihood function. As a solution for this, the log of the likelihood function is used instead, yielding large negative expressions to maximize over. Because this is a monotonic transformation of the original likelihood function, the solution will be the same. The solution will be the absolute value of the negative number that is closest to zero. The log-likelihood function is written as:

Eq. (7)

$$LL = \sum_i^N \log \left[\sum_{c=1}^c P(c: z_i) \prod_{j=1}^J P_{ijk|c}^{r_{ijk}} \right]$$

4.2.2 Marginal Utilities and Marginal Rates of Substitution

It is typical that when attributes of goods are being modeled they have indifference curves that are continuous functions. Analogously, this study has dummy variables with discrete distributions. Simply put, because of the use of dummy variables, indifference curves collapse to individual points of equal utility.

These parameter estimates are generated for all attributes, which are all dummy variables by design. Any given β , such as β_1 , is a scalar that takes on any one value and is the marginal utility of any individual i where U_i is maximized. Therefore, a change in x of one unit ($\Delta x = 1$) will change $V(\cdot)$ by β_1 . By themselves β s tell the sign (i.e., + or -) and significance of the parameter, which has meaning itself.

When comparing parameter estimates, it is beneficial to look for the relative differences between the same MU parameters within each class. This is otherwise known MRS, which is the negative of the ratio for two β s. The MRSs (not MUs) are compared across classes; the biggest relative differences will define the main differences between classes. What is most important is the relative magnitude of the β s, because the utility function only has ordinal properties. For example, the marginal rate of substitution for two marginal utilities, MU_1 and MU_2 can be written as:

Eq. (8)

$$MRS_{12} = \frac{-MU_2}{MU_1}$$

where the MU subscripts (1 and 2) denote two attributes. For example, if attributes 1 and 2 have MRS equal to 1 then those two attributes can be substituted interchangeably with one another to hold utility constant. However, if MRS is greater than 1 ($MRS > 1$) then the attribute in the numerator (MU_2) will dominate the attribute in the denominator. Inversely, if $MRS < 1$ then the importance of the denominator is greater than the attribute in the numerator.

4.2.3 Information Criterion Tables & Class Estimation

Much consideration has been put into evaluating information criteria indices (ICs) in LCMs and selecting the most appropriate and parsimonious indicators for choice experiments. IC is a tool used in many studies to determine model fit and parsimony, and

is often used to determine the number of classes in the underlying data set. Unfortunately, for researchers there is no clear description of how to correctly choose the correct IC for estimation of latent classes.

These parsimony indices acknowledge that while adding more parameters can explain choices better, those additional classes can use up valuable degrees of freedom (df). LGC estimates a few different IC indices, the ones used in this study include: Akaike information criterion (AIC; Akaike, 1987), Bayesian information criterion (BIC; Schwartz, 1978), sample adjusted BIC (SABIC; Sclove, 1987), and corrected Akaike information criterion (CAIC; Bozdogan, 1987).

IC indices are penalized likelihood functions imposed to add weight for each additional parameter and/or sample size for empirical reasons (Yang and Yang, 2007). Generally, the class size with the lowest IC values is the indicator of the number of classes (C). It is typically preferred that the model also produce the largest log-likelihood function using the fewest parameters (Tofighi and Enders, 2008). The uncertainty, however, arises in settling on which IC indicator best represents any given study empirically and technically.

Parsimony indices consider both the likelihood function (Eq. 6) as well as a penalty function for higher numbers of classes. Model parsimony can be used to reduce sampling error as caused by estimation of too many parameters (Dziak et al., 2012). There has not been any consensus on which criterion is best overall, only empirical reasoning and heuristic arguments. That being said, the correct IC is empirical and subjective. This section and Table 4.1 reviews the most significant information criterion indices when using them to select the number of classes in a LCA.

All IC indices discussed in this paper share a commonality in their functional format shown in Eq. (9). All of the underlying IC indices impose a penalty function that adjusts for the number of parameters (p) and increasing sample sizes (N). They also differ by this penalty function imposed on the log-likelihood (LL) function as some typically incorporate heavier penalties on parameter weights and sample size. Most ICs for model selection are derived from a general function where $\log L(\theta_j)$ is the log likelihood obtained from model

j and θ_j is the parameter set of model j . Its specific functional form is as follows (Yang, 2006):

Eq. (9)

$$IC = -2 \log L(\theta_j) + \text{penalty term}(s)$$

Dziak et al. (2012) explain a more general way of modeling the penalized log-likelihood function given:

Eq. (10)

$$-2LL + A_n p$$

where LL is the log-likelihood, A_n is some penalty function of sample size n that varies with the respective IC (see Table 4.1). p is the number of parameters. Both Eq. (9) and (10) find the lowest value of $-2LL + a \text{ penalty term}$.

The above expression is sometimes replaced with $G^2 + A_n p$, where G^2 is $-2LL$ + a function of the unrestricted model (see Yang and Yang, 2007 and Dziak, et al., 2012). The distribution of G^2 is sometimes violated under certain conditions further explained in Appendix H.

Because this regularity condition of the chi-squared statistic being asymptotically distributed is often violated in LCA; additional methods, specifically the bootstrap likelihood ratio test, are used to measure fit and significance of additional classes in the model (see Appendix H).

Table 4.1 shows the different penalty functions for the four IC indices (Dziak et al.,2012).

Table 4.1: Information Criteria LL functions

IC	Penalty Weight
AIC	$A_n = 2$
Sample adjusted BIC (SABIC)	$A_n = \ln \left(\frac{n+2}{24} \right)$
BIC	$A_n = \ln (n)$
CAIC	$A_n = \ln (n + 1)$

4.3 Results

Descriptive statistics of all the sociodemographic variables can be seen in Appendix I. One observation that stands out when taking a closer look at these statistics is that responses to individual political preference contain the most missing data. It was thought that this may be associated with foreign students not having a political preference. However, after further analysis of the data that was proven false. In fact, foreign students showed a high response rate to the political beliefs question. Only one foreign student failed to respond to this specific choice question (see Appendix K). Additionally, the “valid” N found in Appendix I shows a value of 3,946 (valid N= 3,946); this indicates 90% of the sample had no choice data missing.

4.3.1 Selecting the Final Model

Models with 1-5 classes were estimated without (Table 4.2) and with and covariates (Table 4.3) to analyze attribute parameters and various IC scores. All course attributes were tested and were consistently significant. The covariates, however, are a different case. All covariates were tested and incorporated or omitted based upon their significance levels. This ultimately resulted in a final model with three significant covariates discussed in detail in the next section. The number of classes was chosen from Table 4.2, the model without covariates. Although it was not the decisive model, Table 4.3 additionally shows that the SABIC for the 4-class model has the best fit out of the other models.

The model estimates were run multiple times because there are different starting values are used in LGC, and sometimes LGC is susceptible of arriving at a local instead of global max (see Section 4.2.1)¹³ The final model estimates used were the ones that LGC indicated that the “best start seed” was being used in estimation of the underlying model (see LGC 4.0 User’s Guide, 2005 p 39). Based on classic interpretation of IC indices, the lowest value of the IC indicator used indicates the best model. The results for all the IC were different, but most scores, including AIC, AIC3, and SABIC, suggested four latent classes. The SABIC column and 4-class model rows are emphasized in Tables 4.2 and 4.3 shown subsequently. This shows that the bolded 4-class model is superior to the other because it takes on the lowest value (5385.2115) in Table 4.2. Below are the fit indices estimated for models 1-5.

Table 4.2: 1-5 Class Models with No Covariates

Classes	LL	BIC(LL)	AIC(LL)	AIC3(LL)	SABIC(LL)*
1-Class	-2698.18	5475.3605	5420.3596	5432.3596	5437.2570
2-Class	-2658.01	5480.6026	5366.0173	5391.0173	5401.2203
3-Class	-2631.64	5513.4579	5339.2884	5377.2884	5392.7968
4-Class*	-2605.70	5547.1513	5313.3974	5364.3974	5385.2115
5-Class	-2589.38	5600.1026	5306.7644	5370.7644	5396.8839
* Indicates the model chosen and the IC used.					

Table 4.3: 1-5 Class Models with Covariates

Class	LL	BIC(LL)	AIC(LL)	AIC3(LL)	SABIC(LL)*
1-Class	-2520.9144	5120.0053	5065.8287	5077.8287	5081.9041
2-Class	-2477.4179	5143.7624	5012.8357	5041.8357	5051.6846
3-Class	-2437.8831	5175.4429	4967.7661	5013.7661	5029.3885
4-Class*	-2405.3077	5221.0423	4936.6154	4999.6154	5021.0113
5-Class	-2387.1727	5295.5224	4934.3454	5014.3454	5041.5148
* Indicates the model chosen and the IC used.					

¹³ A global max is desired in this study and estimates the best solution out of all possible solutions. A local max is the result of bad starting values.

While the most interesting in terms of parameters, the 4-class model was also the most identifiable by the IC indices, it also has the best parsimony values and most significant parameter estimates for interpretation. See Appendix I for all estimated model statistics. The bootstrap likelihood ratio test (BLRT) was additionally implemented to test for significance when increasing from a three-class to a four- class model. The BLRT is used as a measure of fit of the final model to the data. It was simply used to determine if the increase in the number of classes is significant. The reported p-value of 0.0200 shown in Appendix G shows that it was significant to increase from the three to four class model ($p = 0.0200 \leq \alpha = 0.05$). The BLRT results for an increase from four to five classes informs that additional classes are not improving the model. See Appendix G for an overview of the BLRT.

4.3.2 Interpreting the Final Model

The final 4-class model has three covariates that consistently remained significant: (1) *POLPREF* (political preference) (2) *INTEREST* (economic interest level) and (3) *USA* (domestic or foreign student). The covariate estimates for the significant sociodemographic characteristics can be seen in Table 4.6 and Appendix H The generated p-values of variables are used to indicate each attributes significance. All p-value estimates are conducted on a 95% confidence interval with a significance level of 0.05 ($\alpha = 0.05$). A variable is significant if the critical value is less than 0.05 ($p \leq 0.05$).

The sociodemographic question about individual political preference generated a p-value of 0.0012 ($p = 0.0012$), and resulted in 56.3 % of respondents indicating that they were conservative and 38.2% liberal. The covariate *interest* recorded a p-value of 0.05 ($p = 0.05$). There was a frequency of 61.8% of students that indicated they had a medium level of economic interest and 17% that indicated a high interest. There was a small sample of foreign students (4%). The variable *USA* however, still proved to be significant ($p = 0.021$). Albeit a small proportion of N , the preferences of foreign students should be considered because they make up a large portion of class 2.

The breakdown of the four class sizes for the final model as percentages are as follows. Class 1 has the largest membership (i.e. number of students) with approximately 40% of students. Class 2 membership is slightly higher than Class 3 with 29% of students

falling into Class 2 and roughly 19% for Class 3. Class 4 is the smallest class representing just 12% of the student sample.

In LCMs, as previously described, class preferences are homogeneous “within” class but vary “between” classes. While some classes significantly differ from others, there are also similarities among classes. This does not mean that all attribute preferences are exactly the same, but only that they considered that attribute provided the most utility out of all attributes. The four classes represent four different types of students with distinctly different preferences. Table 4.4 depicts a summary of each classes preferred attribute levels. This table shows what students like most, and in the case of attributes with three levels, what they prefer least.

Model Explanatory Power

LGC produces two types of R^2 estimates, an overall R^2 and a pseudo R^2 . While typical R^2 values indicate the amount of variance explained through an ordinary least squares (OLS) model, there are other considerations when estimating LCMs (see Vermunt and Magidson, 2005 p. 47). For those reasons, other methods can be used to determine the quality of the models explanatory power. Degrees of separation and hit rates are some examples of the techniques used in analysis to essentially replace the use of the R^2 . As an example of tools used to estimate model explanatory power, the hit rate is shown below in Table 4.5.

The hit rate is a technique offer by LGC through analysis of the models ability to correctly predict choices. A “hit” is when the estimated probability of the actual choice is greater than 0.50. As seen in Table 4.5, this model can predict ~79% of responses correctly.¹⁴ The uninformed prior of predicting a pairwise-choice response correctly is 50%, so the final 4-class model definitely incorporates some choice heterogeneity,

¹⁴ The hit rate in Table 4.5 is interpreted follows. There were 1,613 out of 2,018 correct predictions in the model where a respondent chose 1 (i.e., A). There were 1,593 correct predictions of choice 2 (i.e., B). Adding the diagonals where choices were correctly predicted and dividing that sum by the total number of cases predicted (i.e. $1,613 + 1,593 = 3,206 / 4042 = .7931$).

increasing the probability of predicting choice responses by almost 30%. More LGC output can be seen in Appendix D.

Table 4.4: Attribute Preferences by Class

Attribute	Class 1	Class 2	Class 3	Class 4
Focus	Balanced use and Conservation	Resource use/ Energy Markets	Balanced use and Conservation	Conservation/ Sustainability
<i>Least Preferred</i>	<i>Conservation/ Sustainability</i>	<i>Balanced use and Conservation</i>	<i>Resource use/ Energy Markets</i>	<i>Balanced use and Conservation</i>
Credit	Major Requirement Fulfilled	Major Requirement Fulfilled	Major Requirement Fulfilled	Graduation only
Study Hours	Medium (5hrs)	Low (3hrs)	High (7hrs)	Medium (5hrs)
<i>Least Preferred</i>	<i>High (7hrs)</i>	<i>High (7hrs)</i>	<i>Medium (5hrs)</i>	<i>Low (3hrs)</i>
Professor	Tenured	Tenured	Instructor/Lecturer	Tenured
<i>Least Preferred</i>	<i>Junior Professor</i>	<i>Instructor/Lecturer</i>	<i>Tenured</i>	<i>Junior Professor</i>
Paper Req.	Small	Large	Small	Small
Time	Midday/ Morning	Late Afternoon	Morning	Midday
<i>Least Preferred</i>	<i>Late Afternoon</i>	<i>Morning</i>	<i>Late Afternoon</i>	<i>Morning</i>
Day	MWF	Tu/Th	Tu/Th	Tu/Th
Semester	Spring	Spring	Fall	Fall

Table 4.5: Model Hit Rate

Hit-Rate Prediction Table			
Observed	1	2	Total
1	*0.80	0.21	2044
2	0.20	*0.79	1998
Total	2018	2024	4042
* Indicates (in percentages) the number of choices the model correctly predicts.			

4.3.3 Model Estimation

One-Class Model Estimates

Using LGC, many statistics were generated that give us estimates of parameters for a one-class model. This is essentially a DCM before choice heterogeneity is allowed through multiple classes and covariates (i.e., LCA). Assuming only one class ($c = 1$) forces everyone to have the same preferences, which may produce biased results. This is because in reality, the model is highly non-linear and everyone does not have homogenous preferences. This estimation gives us *ceteris paribus* parameter estimates of a homogenous model which can be compared with to final model once heterogeneity is incorporated through latent class membership. Parameter estimates for the one-class model can be seen in Table 4.6 below, alongside the final 4-class parameter estimates. This simple, one-class model is estimated to investigate whether class-specific preference parameters vary greatly from the homogenous model parameters. Table 4.6 reveals that they do. Lastly, in Table 4.6 there are no covariate estimates for the sociodemographic questions, particularly because the 1-class model because covariates are only used to explain membership in more than one class.

Final LCM Estimates

When interpreting the parameter estimates (in general, not just for one-class models), the largest value indicates the level of the respective attribute that has the most impact on utility for the population as a whole. Table 4.6 below has been bolded to emphasize the most preferred level for the attribute *focus* for all four classes. For example, resource use/energy markets (1.2022) is in bold to show “Class 2” most prefers this level of course focus over all the others. This is the most preferred course focus; that is, it yields the highest utility of the three attribute levels for *focus*. The zeros for various attribute levels respectively indicates that attribute level is the base-case. While these attributes are omitted from the final utility function because they can be estimated by incorporating the other level(s) of the respective attribute, they are shown in Table 4.6 for ease of interpretation of parameter estimates. Furthermore, the model can only be estimated with these parameters.

Table 4.6 shows all parameter estimates (i.e., MUs) for all attributes, and uniquely large parameters (in absolute value) are bolded for ease of interpretation. The bolded attribute levels are the “most preferred” class preferences shown in Table 4.4. The “least preferred” are those with the lowest estimates.

It is assumed that in this experiment V_{ijk} is a function of the following dummy variables: the focus of the course, whether or not it is required course, the number of weekly study hours, the rank of the professor, the size of the term paper, the time of day of the course, the day of the week of the course, and the semester the course is offered in. Students economic interest, whether the student is domestic or foreign, and the students’ political preference are significant but are not included in the utility function. As explained in Section 3.1.1, these attributes break into multiple dummy variable, thus, the utility function for individual i can be written as:

Eq. (11)

$$V_{ij} = \beta_{1ij}RES.USE_{ij} + \beta_{2ij}CONSERV_{ij} + \beta_{3ij}GRAD_{ij} + \beta_{4ij}HIGH_{ij} + \beta_{5ij}MED_{ij} + \beta_{6ij}INSRUCT_{ij} + \beta_{7ij}JUN.PROF_{ij} + \beta_{8ij}LARGE_{ij} + \beta_{9ij}MORNING_{ij} + \beta_{10ij}MIDDAY_{ij} + \beta_{11ij}TU.TH_{ij} + \beta_{12ij}SPRING_{ij}$$

This table consists of parameter estimates from the one-class model, and the final 4-class model. In comparison to the homogenous model (1-class), one can intuitively conclude that the heterogeneous model produces four preference classes that are distinctly different in membership size and parameter estimates, and all greatly differ from the one-class model. Additionally, the p-values shown above for each attribute are significant in the 4-class model, a large improvement from the 1-class model. See Section 4.4 for a description of each class’s preferences based on the Table 4.6 parameters.

Table 4.6: Attribute Parameters

Attributes	One-Class Model	p-value	Class1	Class2	Class3	Class4	p-value
<i>focus</i>		0.24					0.0017
RESOURCE USE/ ENERGY MARKETS	0.2336		-2.0571	1.2022	-3.1962	7.809	
CONSERVATION/ SUSTAINABILITY	0.2242		-2.3994	0.5638	-1.891	12.8468	
BALANCED USE AND CONSERVATION	0		0	0	0	0	
<i>major</i>		2.20E-06					0.026
GRADUATION ONLY	-0.2607		-1.2961	-0.3836	-0.9671	2.9897	
MAJOR REQUIREMENT FULFILLED	0		0	0	0	0	
<i>studyhr</i>		6.30E-30					7.40E-07
HIGH (7HRS)	-0.3054		-1.6259	-0.1812	0.2068	3.9139	
MED (5HRS)	0.4487		1.9956	-0.0127	-0.0849	5.8404	
LOW (3HRS)	0		0	0	0	0	
<i>profrank</i>		7.20E-09					0.00014
INSRUCTOR/LECTURER	-0.2776		-0.3689	-0.2581	3.1144	-6.4066	
JUNIOR PROFESSOR	-0.4903		-0.9871	-0.0879	2.4118	-9.681	
TENURED PROFESSOR	0		0	0	0	0	
<i>smpaper</i>		7.80E-27					2.80E-08
LARGE % OF GRADE	-0.6042		-2.326	0.0509	-3.3061	-0.3661	
SMALL % OF GRADE	0		0	0	0	0	
<i>time</i>		1.30E-07					5.90E-05
MORNING	-0.1254		0.0566	-1.2057	3.7481	-2.2075	
MIDDAY	0.2603		0.0538	-0.0976	2.3076	2.4941	
LATE AFTERNOON	0		0	0	0	0	
<i>MWF</i>		0.14					0.00024
T/TH	0.1076		-1.607	0.2912	0.9151	4.5988	
MWF	0		0	0	0	0	
<i>fall</i>		0.24					0.04
SPRING	0.0528		0.6093	0.2862	-3.1399	-3.6401	
FALL	0		0	0	0	0	

4.3.4 Student Tradeoffs

It is important to note that large negative parameters of opposite signs but similar absolute values can be used to offset attributes that make students “worse off” for attributes that students actually prefer. Simply put, changing one course attribute can have a positive or negative impact on students’ decision to enroll in any underlying course. This section demonstrates how to interpret different student tradeoffs between different course attributes by using MUs to estimate MRSs.

For example, the bolded fields in Table 4.7 show that for Class 2 only ($c = 2$) a course offered in the morning (-1.2057), which is not preferred, can be potentially offset with a course that focuses on resource use and energy markets (1.2022). This is an example of the MRS between two attribute parameters in one specific class, explained in Eq. (8) (section 4.2.2). In other words, Class 2 may be willing to attend a morning course if that course focus is their most preferred; in the case of this study, that course preference is resource use and energy markets.

Table 4.7: Tradeoff between Class Focus and Class Time

Attributes	Class2
<i>focus</i>	
RESOURCE USE/ENERGY MARKETS	1.2022
CONSERVATION/SUSTAINABILITY	0.5638
BALANCED USE AND CONSERVATION	0
<i>time</i>	
MORNING	-1.2057
MIDDAY	-0.0976
LATE AFTERNOON	0

Table 4.7 shows how easily attributes with MRS equal to 1 ($MRS = 1$) can tradeoff. There are however other attribute tradeoffs with unique parameters. One example of student attributes that caused a significant change in students’ probability of enrolling in a course was the changing of course professor from instructor/lecturer to a tenured professor.

All else constant there was an 22% decrease, for Class 1, in the probability of enrolling when the professor is not a Tenured professor. Class 2 probability of enrolling is negatively affected by the change to instructor/lecturer by 20%. Class 3 oddly prefers an instructor/lecturer over all professor types, so their probability actually increases. The theory as to why this may be can be explained by Appendix B. Class 4 only decreases by 1% when this professor change is made. However, the choice question sampled here includes conservation/sustainability as an attribute in alternative set A, and Class 4 largely prefers this course focus to any other. So much that the type of professor teaching course, nor any other course attribute, could produce enough utility to be a viable tradeoff for members of Class 4. In summary, there is positive increase of the probability of enrollment for 81% of the sample if there is a tenured professor teaching the course.

Further analysis shows that the probability of enrollment increases for 71% of students if the size of the term paper is changed from large to small. Additionally, there is a negative effect on 71% of the sample when the time of day of the course is changed from morning to late afternoon. Class 3 is, however, affected the most significantly.

4.4 Class Descriptions

Class 1

1. Class 1 constitutes approximately 40% of the population.
2. This class consists of almost 70% conservative students with a medium interest level in economics.
3. Being the largest class, Class 1 consists of some foreign students (~2%), however the dominating majority of students in this class are domestic.
4. In terms of attributes, they prefer courses that provide credit for their majors and focuses on balancing use and conservation. Class 1 notably takes a strong disliking towards the other two levels of course focus in favor of balance use and conservation.
5. They prefer that classes are taught by a tenured professor.
6. A medium level of homework hours is preferred (5 hours).
7. Their preferences are consistent with courses offered on MWF during the spring semester.

8. If a research paper is required, these students prefer that it counts for a smaller portion of their final grade.
9. Class 1 is indifferent to class time with MU roughly equal when choosing between morning and midday courses. Late Afternoon courses are the least preferred by Class 1.

Class 2

1. Class 2 represents 29% of the student population.
2. This class is about 60% conservative.
3. Unlike Class 1, Class 2 prefers courses focused towards resource use/energy markets.
4. Class 2 is the sole class with a “high” level of interest in economics
5. They prefer a low amount of study hours and that their courses are also taught by a tenured professor.
6. This class also contains the most foreign students of all classes which is consistent with expectation that foreign students would prefer resource use and energy courses, and are often in the master’s program due to mining interests. Foreign student presence in other classes is extremely minimal.
7. It is also preferred by class 2 that a course contributes credit towards their majors and is offered on Tu/Th.
8. Like class 1, a spring semester enrollment is preferred over fall.
9. This class uniquely prefers a large research paper requirement and courses scheduled in late afternoon.
10. Class 2 least prefers courses on balanced use and conservation and courses taught by an instructor/lecturer.

Class 3

1. Class 3 represents 19% of the population.
2. This class consists of primarily conservative individuals in terms of political preference. Foreign students are nearly non-existent in this class.
3. This class has a medium interest in economics.

4. Like class 1 this class has high preference for courses with a balanced use and conservation focus.
5. Class 3, like the preceding two classes, prefer that the course counts towards major credit.
6. These individuals additionally have preference for a small research requirement.
7. Unique from all other classes they desire a high level of study hours and for the course to be taught by an instructor/lecturer.
8. They prefer Tu/Th courses and fall semester enrollment.
9. They uniquely have a high preference that their courses are taught in the morning.
10. This class least prefers that the course focuses on resource use and energy markets versus all other options.

Class 4

1. Class 4 represents the smallest group of students at 12%.
2. This class is almost entirely made up of liberals with a medium level of economic interest. Foreign students are nearly non-existent in this class.
3. Results show this class highly prefers courses that focus on conservation/sustainability, and they additionally show a strong preference for courses on resource use and energy markets, however, the clear frontrunner is conservation/sustainability.
4. Class 4 is the unique class where the individuals would not mind at all taking a class that will not contribute credit towards their academic majors.
5. Like class 1, they prefer a medium level of study hours and a small research paper requirement.
6. Class 4 highly prefers their courses to be taught on Tu/Th and during the fall semester.
7. They highly prefer a tenured professor exhibiting extreme negative parameters for both junior professors and instructors/lecturers.

Class Summary

Preference for balanced use and conservation courses superseded all other course focus attribute levels by being the dominant focus for class 1 and 3. Combining those two classes, roughly 60% of students prefer a balance use and conservation course focus. Additionally, the makeup of the latter 60% consists of conservative students by large majority.

It is worth noting that no class prefers to be taught by a junior (assistant) professor. Classes 1, 2, and 4 prefer a more seasoned tenured professor rather than instructor/lecturer or junior professor. Over all four classes, tenured professors are preferred over the other two alternative teacher ranks. Class 3 prefers an instructor/ lecturer. Furthermore, classes 1 and 2 prefer spring semester enrollment and the smaller classes, 3 and 4, prefer fall semester enrollment.

The majority (classes 1-3) have similar preferences for courses that provide credit towards their majors. The preferred amount of study hours varies amongst all classes; however, the majority of students prefer a medium level of weekly study hours (5 hrs). Every class had a different preferred course time. Class 1 was the only class that chose MWF as preferred days for courses, whereas classes 2 through 4 chose Tu/Th. Class 1 is however the largest class so their MWF preference is still a significant attribute level to consider even though the majority preference across the entire sample is Tu/Th. Classes 1 through 3 tend to prefer credit allocation towards their major. Diverging is class 4 where students, whom are mostly liberal, highly prefer elective courses outside their major. The foreign students appear to be predominantly in class 2 with a small influence on class 1. Classes 3 and 4 contain almost no foreign students.

4.5 Student Awareness Analysis

A correlation matrix (See Appendix F) was developed to incorporate the first set of three Likert-scale survey questions into the analysis. The purpose of this analysis is to investigate the relationship between the awareness of the students surveyed (questions 1 and 2) and their likelihood of enrolling in one or more of the environmental courses (question 3).

All correlation coefficients had a positive relationship, and were significant at the 0.01 level in a 2-tailed t-test using the Spearman's rho correlation matrix. The relationship between students' awareness of NR courses offered and awareness of the NRE master's program asked in questions 1 and 2 was the strongest relationship. This was expected to be the stronger of the ratios because naturally students who don't know of the undergraduate course offerings most likely would not be informed on the NRE master's program either. The likelihood of enrollment and awareness of NR economics courses is slightly stronger than that of awareness of the master's program. Nonetheless, the relationship between students' likelihood of enrolling and awareness were significant at a high confidence level (99%; see Appendix F).

Frequency tables (Appendix J) are used to analyze questions 1-3 to explore the repeat occurrence of each level of awareness for respondents prior to taking the survey. Disappointingly, ~42% of all students stated that they were "not at all aware" of the NRE courses offered at Michigan Tech. An additional 26% of students responded that they were "a little aware" of the underlying courses offered. This leaves 68% of students that have little to no idea of NRE courses offered at the university.

64.5% of all students who took the survey stated that they were "not at all aware" of the master's program, which can be expected when nearly half of the sample states that they are completely unaware of any of the NR courses offered. A total of 9% of respondents stated that they were "extremely likely" to enroll in courses discussed in questions 1 and 2.

Furthermore, crosstab analysis (Appendix K) reveals that "zero" foreign students were "extremely aware" of the NRE master's program. Additionally, students with low economic interest have extremely low (1%) likelihood of enrollment, whereas students with high economic interest show a 31% likelihood to enroll in one of the underlying course options.

Chapter 5: Discussion

The study brought to light several important discussions. In order to maximize student welfare, significant benefits may come at the expense of faculty desires. In the pursuit of optimizing electives and increasing course enrollment, it is suggested a series of important lecture and administrative applications based off of the present results should be considered by university programs seeking to do the following: (1) increase the quality and relevance of electives offered by understanding what attributes students prefer more or less when choosing among alternatives; and (2) explore how to use limited marketing and student outreach financial resources to target students that are most likely to enter and complete courses and programs. While some of the recommendations may already be used by MTU in other areas of study, this discussion provides recommendations on how to attract students in each of the estimated four classes, to NRE courses (The surveyed courses that the sample was drawn from can be seen in Appendix L). This is followed by general recommendations to consider for electives as a whole.

5.1 Attracting Class-specific Individuals

Class 1

To attract students in Class 1 (largest class size), universities and more specifically professors should do the following:

1. Increase their outreach to conservative student groups. Some examples include: ROTC programs, business students, campus Republicans, and students with “some” economic interest (i.e. students currently enrolled in economics courses). A few outreach techniques include mass emails, campus-wide brochures (specifically in buildings where conservative students spend the most time). This study found the most conservative students to be STEM (Science, Technology, Engineering, and Mathematic) majors.

Class 2

1. The recommendation for the preceding class, Class 1, is applicable to Class 2 as well on the grounds of a high percentage of conservative students.

2. This class has the highest economic interest so they are likely already enrolled in more than one economic course. Reaching these students for purposes of this study should be the easiest because of their high involvement with economic professors. Targeting specific clubs and organizations such as the economic and business clubs is can assist in finding these types of students.

Class 3

1. Students in Class 3 interestingly prefer an instructor/lecturer over all other teacher rankings. It is not clear as to why, however, Appendix B assists in explaining eccentric responses. Conversely, studies support opposing views on the quality of academics taught by non-tenured professors (Kezar and Maxey, 2013). However, the dominating 82% of all students' clear preference for a tenured professor should outweigh any drastic changes to instructor type in favor of Class 3. With exception of Class 3, tenured professors tend to have a large positive influence on students' willingness to enroll in the underlying course options. This is in stark contrast to the modern movement in academia to hire more lecturers and instructors with less job security.
2. Class 3 additionally prefers classes with light research requirements; however, they prefer a high amount of study hours. There is a clear tradeoff between the two attributes, even though the respective attribute parameters are not close in comparison. Class 3, nonetheless, would clearly rather study than write research papers, suggesting for teachers to offer an alternative option between the two main elements of curriculum schedules.

Class 4

1. Students in class 4 belong to the smallest membership class. It would be logical to suggest that no recommendation that contributes negatively to the other populations be implemented in favor of Class 4, because it represents the lowest number of students. However, this class's extremely high preference for conservation courses of study makes them relevant regardless of the fact they represent 12%.

With troublesome conservation issues (e.g., climate change) affecting everyday lives, it is better to be inclusive rather than exclusive in an area that focuses on vital environmental issues. The latter, along with being the only class that prefers graduation - only courses over major-fulfilling courses makes it hard for a NRE program to ignore the genuine interest in the field; especially since liberal students tend to be the most concerned with a healthy environment and are abundant in class 4. In layman's terms, Class 4 is the only class who would prefer to enroll in the conservation or sustainability focused courses. As a bonus, their choice of taking NRE courses as electives indicates they do not mind the least bit that they may not receive credit towards their major. That being said Class 4 may be worthy of special attention from faculty or administrators in order to attract this group of students. For example, offering courses that focus on conservation and sustainability midday on Tu/Th is desired by Class 4.

5.2 Study Recommendations

The first suggestion for MTU to consider concerns the fact that courses focuses on balanced use and conservation are highly preferred to other alternatives by primarily conservative students, who make up the majority of students in the sample. The program should focus on targeting these students by simply offering more balance and conservation courses that are offered in spring, midday. The majority of these students who prefer balance use and conservation (59%) prefer these classes on MWF.

Secondly, overall there is a significant preference for tenured professors over the alternatives instructor/lecturer and junior professor. This states that students prefer a more experienced professor opposed to a non-tenured junior professor or lecturer. Logically, it would make sense to that a tenured professor is preferred because of some inconsistencies with lecturers and the typical strenuous research expectations junior professors endure when attempting to reach tenure, which might compromise teaching quality (see Kezar and Maxey, 2013).

The third suggestion surrounds the weekly study hours required by courses. It is clear that all classes except one (Class 2) prefer medium study hours. This is inconsistent with the hypothesis that most students would choose low study hours. This may be case specific being that MTU students are faced with a rigorous curriculum. Easy classes will

not attract students who “want their money’s worth.” It is interesting to discover that students do not prefer low study hours; this suggests that students prefer to do “some” work, signaling a desire to learn.

Next, are 4 additional general suggestions that can be adopted by MTU and other comparable universities looking to reach a similar objective.

1. It is recommended that the university schedule “required” courses early in the morning. While the distaste for morning courses was unanimous, if required/major-specific courses are scheduled in the mornings, then students have no choice but to rise early and take their required courses. This allows for schedules to be open for desired electives midday, which is preferred by all, and could increase enrollment due to strategic scheduling.
2. A t-shaped curriculum should be structured. More specifically, there should be classes with highly specific and closely related topics. Conversely, there should also be offerings of “survey courses” that offer a cursory discussion of a broad variety of loosely related topics.
3. This study shows that students are discouraged by courses with large research paper requirements. Since it is a requirement of many courses in the NRE program offered by MTU, professors might consider an alternative to a large research assignment. It is understood research is an important area in the underlying field so one suggestion may be to break the research requirement into a few smaller research assignments that are separate or together make up a complete research paper.
4. Touching on a more sensitive area, the classic organization of classes around faculty schedules may be driving down enrollment. This is supported by the fact that most students dislike the current scheduling of NRE courses. Professors, understandably, often look to clump their classes on the same days during a certain time frame for various efficiency reasons. If faculty could finesse their schedules to teach for five days a week (M-F) then it is expected to increase likelihood to enrollment.

5.3 Increasing Student Awareness

It is important that measures are taken in order to increase the general awareness, or knowledge, of students with respect to NRE studies. Based on the survey “awareness” questions 1 through 3 it is suggested that not many students are informed about the NRE master’s program or undergraduate course offerings. A logical solution could be to increase promotion of the program and NRE courses. That may itself have a positive impact on enrollment numbers. For example, require that every economics professor introduce and be an advocate for the alternative economic avenues uniquely offered by the university (e.g., NRE master’s program) in their introductory courses. This is a cost-free strategy that could be implemented by dozens of professors, and is also a direct marketing method that helps enhance the students’ engagement. The results suggest that more students might be interested in the underlying programs if they simply knew about them.

Limited resources can be used to increase enrollment by targeting the key students that are (1) unaware of NR courses and (2) are, at minimum, “somewhat likely” to enroll in economics courses. Targeting students below the “somewhat likely” threshold produces inefficient results, and can in result be a waste of university budgeting funds. Furthermore, the correlation between likelihood of enrollment and student economic interest proved significant at a 99% confidence interval; $p \geq 0.01$ (see Appendix F).

Given the opportunity, targeting the students with higher likelihood of enrollment may inadvertently result in the captivation of some students with low economic interest or even respondents who are “a little likely” to enroll in NR economics courses or the master’s program. The same empirical theory, however, may not be applicable in the inverse situation. Creating marketing materials or campaigns for broad audiences, and assuming that the interested students will absorb the information and enroll, has potential benefits, but at a significant cost. One risk is that the students most likely to enroll could simply not see or engage in the underlying marketing medium. Expressed in the previous class descriptions and recommendations is the characteristics of the students or classes that should and should not be targeted. For example, more efforts should be focused on students actually interested and resources should not be exhausted on students with low economic interest. In the interest of efficiency, the majority of resources should be concentrated on

students that might actually enroll. Albeit, if the funds are available to target the entire student population then it should be done, because absent all other analysis, the majority of students are not aware of NR economics courses nor informed of the NRE master's program.

Additional Recommendations

It is obvious that technology proliferation has significantly influenced the world exponentially. As a result, industries have been disrupted causing decreases in demand for human labor. While institutional teaching has not been abolished, there are increasing challenges to capture students due to the competing online forces (e.g., Khan Academy, YouTube, and online universities and academic institutions). Mentioned before, Apperson et al. (2008) concluded that students preferred PowerPoint taught courses. This asserts students' growing preference to learn using online applications, not to mention technology has advanced significantly since the latter study.

In fact, the recent emergence of Massively Open Online Courses (MOOCs) at high-profile universities such as Harvard has taken academia by storm. MOOCs are a newer form of distance education that offers open online courses for thousands of students at a time. They are higher education courses which range over a broad range of topics and concentrations and act as a substitute to traditional academic institutions (see Fischer, 2014). Furthermore, a survey conducted by *Insider Higher Ed* supports the notion MOOCs are consistent with the trend at all school levels to replace tenured professors in specialized fields with lower-salaried instructors without tenure or doctorates; but who are flexibly trained to teach a variety of large entry-to-mid level courses (Jaschik and Lederman, 2014). With respect to budget constraints, these instructors can also be added to or removed from faculties as budgets allow. This indicates a trend in course offerings moving away from specialized courses with small enrollment. This could be a problem because it is these courses that provide job diversity and students with unique skill sets. Professors should use modern technology in their favor to enhance student learning and additionally conform to student preferences that are within reason.

Future Studies

Acknowledging that there may be opportunities for improvement, this section discusses known areas that could be altered in future studies. One improvement to be made is in the potential ambiguity with attribute terms like conservation/sustainability in the survey. There was, however, limited space and time, and conservation versus resource markets and energy use resonated with liberal and conservative students, respectively, as expected. Additionally, the pretests demonstrated strongly that survey respondents understood the terminology and the choice tasks. Nonetheless, a description or key explaining the variables further could be beneficial in future studies due to feedback relating to the potential misunderstanding of some terms.

References

- Adamowicz, W. L. (2004). What's it worth? An examination of historical trends and future directions in environmental valuation. *Australian Journal of Agricultural and Resource Economics*, 48(3), 419-443.
- Akaike, H. (1987). Factor analysis and AIC. *Psychometrika*, 52(3), 317-332.
- Alpizar, F., Carlsson, F., & Martinsson, P. (2003). Using choice experiments for non-market valuation. *ECONOMIC ISSUES-STOKE ON TRENT*-, 8(1), 83-110.
- Apperson, J. M., Laws, E. L., & Scepansky, J. A. (2008). An assessment of student preferences for PowerPoint presentation structure in undergraduate courses. *Computers & Education*, 50(1), 148-153.
- Arcidiacono, P., & Jones, J. B. (2003). Finite mixture distributions, sequential likelihood and the EM algorithm. *Econometrica*, 71(3), 933-946.
- Arcidiacono, P., Bayer, P., Blevins, J. R., & Ellickson, P. B. (2016). Estimation of dynamic discrete choice models in continuous time with an application to retail competition. *The Review of Economic Studies*, rdw012.
- Bateman, I., Day, B., Loomes, G., & Sugden, R. (2006). *Ranking versus choice in the elicitation of preferences*. Working Paper, available at http://www.uea.ac.uk/eo/people/add_files/loomes.
- Behrman, J., Tincani, M., Todd, P. E., & Wolpin, K. I. (2014). The impact of school voucher systems on teacher quality in public and private schools: The case of Chile. mimeo, U. of Pennsylvania.
- Ben-Akiva, M., & Boccara, B. (1995). Discrete choice models with latent choice sets. *International Journal of Research in Marketing*, 12(1), 9-24.
- Ben-Akiva, M., Meersman, H., & Van de Voorde, E. (2013). Recent developments in freight transport modelling. *Freight Transport Modelling*, 1.
- Birol, E., Karousakis, K., & Koundouri, P. (2006). Using a choice experiment to account for preference heterogeneity in wetland attributes: the case of Cheimaditida wetland in Greece. *Ecological economics*, 60(1), 145-156.
- Boxall, P. C., & Adamowicz, W. L. (2002). Understanding heterogeneous preferences in random utility models: a latent class approach. *Environmental and resource economics*, 23(4), 421-446.

- Bozdogan, H. (1987). Model selection and Akaike's information criterion (AIC): The general theory and its analytical extensions. *Psychometrika*, 52(3), 345-370.
- Breffle, W. (2008). In pursuit of optimal design: a guide for choice experiment practitioners. *International Journal of Ecological Economics and Statistics*, 14.
- Breffle, W. S., & Maroney, K. K. (2009). The restoration of fishing services and the conveyance of risk information in the Southern California Bight. *Marine Policy*, 33(4), 561-570.
- Briesch, R. A., Chintagunta, P. K., & Matzkin, R. L. (2012). Nonparametric discrete choice models with unobserved heterogeneity. *Journal of Business & Economic Statistics*.
- Brock, W. A., & Durlauf, S. N. (2001). Discrete choice with social interactions. *The Review of Economic Studies*, 68(2), 235-260.
- Bunch, D. S., Louviere, J. J., & Anderson, D. (1996). A comparison of experimental design strategies for multinomial logit models: The case of generic attributes. *University of California Davis Graduate School of Management Working Paper*, 11-96.
- Campbell, D., Hutchinson, G., & Scarpa, R. (2005, March). Using choice experiments to value farm landscape improvements: implications of inconsistent preferences. In *Applied Environmental Economics Conference, London*.
- Campbell, D., Hutchinson, W. G., & Scarpa, R. (2006). Using discrete choice experiments to derive individual-specific WTP estimates for landscape improvements under agri-environmental schemes: evidence from the rural environment protection scheme in Ireland. *Working Paper No. 26.2006*.
- Campbell, D., Hutchinson, W. G., & Scarpa, R. (2008). Incorporating discontinuous preferences into the analysis of discrete choice experiments. *Environmental and resource economics*, 41(3), 401-417.
- Campbell, D., Hensher, D. A., & Scarpa, R. (2011). Non-attendance to attributes in environmental choice analysis: a latent class specification. *Journal of environmental planning and management*, 54(8), 1061-1076.
- Carlsson, F., Frykblom, P., & Liljenstolpe, C. (2003). Valuing wetland attributes: an application of choice experiments. *Ecological Economics*, 47(1), 95-103.
- Champ, P. A., Alberini, A., & Correias, I. (2005). Using contingent valuation to value a noxious weeds control program: the effects of including an unsure response category. *Ecological Economics*, 55(1), 47-60.

- Chan, C. Y., De Peretti, C., Qiao, Z., & Wong, W. K. (2012). Empirical test of the efficiency of the UK covered warrants market: Stochastic dominance and likelihood ratio test approach. *Journal of Empirical Finance*, 19(1), 162-174.
- Chestnut, L. G., Rowe, R. D., & Breffle, W. S. (2012). Economic Valuation of Mortality-Risk Reduction: Stated Preference Estimates from the United States and Canada. *Contemporary Economic Policy*, 30(3), 399-416.
- Clark, M. D., Determann, D., Petrou, S., Moro, D., & de Bekker-Grob, E. W. (2014). Discrete choice experiments in health economics: a review of the literature. *Pharmacoeconomics*, 32(9), 883-902.
- Cunningham, T. R., & Geller, E. S. (2011). What do healthcare managers do after a mistake? Improving responses to medical errors with organizational behavior management. *Journal of Communication in Healthcare*, 4(2), 70-87.
- Dayton, C. M., & Macready, G. B. (1988). Concomitant-variable latent-class models. *Journal of the American Statistical Association*, 83(401), 173-178.
- Daziano, R. A., Miranda-Moreno, L., & Heydari, S. (2013). Computational Bayesian statistics in transportation modeling: from road safety analysis to discrete choice. *Transport reviews*, 33(5), 570-592.
- Dziak, J. J., Coffman, D. L., Lanza, S. T., & Li, R. (2012). Sensitivity and specificity of information criteria. *The Methodology Center and Department of Statistics, Penn State, The Pennsylvania State University*, 1-10.
- Dziak, J. J., Lanza, S. T., & Tan, X. (2014). Effect size, statistical power, and sample size requirements for the bootstrap likelihood ratio test in latent class analysis. *Structural equation modeling: a multidisciplinary journal*, 21(4), 534-552.
- Fischer, G. (2014). Beyond hype and underestimation: identifying research challenges for the future of MOOCs. *Distance education*, 35(2), 149-158.
- Gibson, J. M., Rigby, D., Polya, D. A., & Russell, N. (2015). Discrete choice experiments in developing countries: willingness to pay versus willingness to work. *Environmental and Resource Economics*, 1-25.
- Greene, W. H., & Hensher, D. A. (2003). A latent class model for discrete choice analysis: contrasts with mixed logit. *Transportation Research Part B: Methodological*, 37(8), 681-698.
- Green Jr, B. F. (1951). A general solution for the latent class model of latent structure analysis. *Psychometrika*, 16(2), 151-166.

- Hackbarth, A., & Madlener, R. (2013). Consumer preferences for alternative fuel vehicles: A discrete choice analysis. *Transportation Research Part D: Transport and Environment*, 25, 5-17.
- Hanley, N., Mourato, S., & Wright, R. E. (2001). Choice Modelling Approaches: A Superior Alternative for Environmental Valuations. *Journal of economic surveys*, 15(3), 435-462.
- Hill, M., Fisher, J., Chitty, L. S., & Morris, S. (2012). Women's and health professionals' preferences for prenatal tests for Down syndrome: a discrete choice experiment to contrast noninvasive prenatal diagnosis with current invasive tests. *Genetics in Medicine*, 14(11), 905-913.
- Holt, J. A., & Macready, G. B. (1989). A simulation study of the difference chi-square statistic for comparing latent class models under violation of regularity conditions. *Applied Psychological Measurement*, 13(3), 221-231.
- Hoyos, D., Mariel, P., & Hess, S. (2015). Incorporating environmental attitudes in discrete choice models: An exploration of the utility of the awareness of consequences scale. *Science of the Total Environment*, 505, 1100-1111.
- Huber, J., & Zwerina, K. (1996). The importance of utility balance in efficient choice designs. *Journal of Marketing research*, 307-317.
- Jaschik, S. & Lederman, D. (2014) The 2014 Inside Higher Ed survey of college & university business officers. Retrieved from Inside Higher Ed website: <https://www.insidehighered.com/booklet/survey-college-and-university-business-officers>
- Johnson, F. R., Lancsar, E., Marshall, D., Kilambi, V., Mühlbacher, A., Regier, D. A., Bresnahan, B., Kanninen, W.W., & Bridges, J. F. (2013). Constructing experimental designs for discrete-choice experiments: report of the ISPOR conjoint analysis experimental design good research practices task force. *Value in Health*, 16(1), 3-13.
- Jung, T., & Wickrama, K. A. S. (2008). An introduction to latent class growth analysis and growth mixture modeling. *Social and Personality Psychology Compass*, 2(1), 302-317.
- Kim, J., Allenby, G. M., & Rossi, P. E. (2002). Modeling consumer demand for variety. *Marketing Science*, 21(3), 229-250.
- Kezar, A., & Maxey, D. (2013). The Changing Academic Workforce. *Trusteeship*, 21(3), 15-21.
- Kuhfeld, W. F. (2010). Conjoint Analysis. *SAS Tech. Pap., MR H, 2010*, 681-801.

Lancsar, E., & Louviere, J. (2006). Deleting 'irrational' responses from discrete choice experiments: a case of investigating or imposing preferences. *Health economics*, 15(8), 797-811.

Lancsar, E., Louviere, J., Donaldson, C., Currie, G., & Burgess, L. (2013). Best worst discrete choice experiments in health: Methods and an application. *Social science & medicine*, 76, 74-82.

Lanza, S. T., & Rhoades, B. L. (2013). Latent class analysis: an alternative perspective on subgroup analysis in prevention and treatment. *Prevention Science*, 14(2), 157-168.

Lockshin, L., Jarvis, W., d'Hauteville, F., & Perrouy, J. P. (2006). Using simulations from discrete choice experiments to measure consumer sensitivity to brand, region, price, and awards in wine choice. *Food quality and preference*, 17(3), 166-178.

Louviere, J. J., Eagle, T. C., & Cohen, S. H. (2005). Conjoint analysis: Methods, myths and much more. *Documento de trabajo del Censoc*, (05-001), 31.

Louviere, J. J. (2006). What you don't know might hurt you: some unresolved issues in the design and analysis of discrete choice experiments. *Environmental and Resource Economics*, 34(1), 173-188.

Louviere, J. J., & Lancsar, E. (2009). Choice experiments in health: the good, the bad, the ugly and toward a brighter future. *Health Economics, Policy and Law*, 4(04), 527-546.

Louviere, J. J., Flynn, T. N., & Carson, R. T. (2010). Discrete choice experiments are not conjoint analysis. *Journal of Choice Modelling*, 3(3), 57-72.

Lusk, J. L., & Norwood, F. B. (2005). Effect of experimental design on choice-based conjoint valuation estimates. *American Journal of Agricultural Economics*, 87(3), 771-785.

Maringe, F. (2006). University and course choice: Implications for positioning, recruitment and marketing. *International Journal of Educational Management*, 20(6), 466-479.

McCorskey, J. C., & McVetta, R. W. (1978). Classroom seating arrangements: Instructional communication theory versus student preferences. *Communication Education*, 27(2), 99-111.

McFadden, D. (1973). Conditional logit analysis of qualitative choice behavior.

McFadden, D. (1994). Contingent valuation and social choice. *American Journal of Agricultural Economics*, 76(4), 689-708.

- Michaud, C., Llerena, D., & Joly, I. (2012). Willingness to pay for environmental attributes of non-food agricultural products: a real choice experiment. *European Review of Agricultural Economics*, jbs025.
- Morey, E., Thacher, J., & Breffle, W. (2006). Using angler characteristics and attitudinal data to identify environmental preference classes: a latent-class model. *Environmental and Resource Economics*, 34(1), 91-115.
- Morey, E. R., & Thacher, J. A. (2012). Using choice experiments and latent-class modeling to investigate and estimate how academic economists value and trade off the attributes of academic positions. *Working Paper, available at <http://www.colorado.edu/economics/morey/discuss.html>*.
- Nylund, K. L., Asparouhov, T., & Muthén, B. O. (2007). Deciding on the number of classes in latent class analysis and growth mixture modeling: A Monte Carlo simulation study. *Structural equation modeling*, 14(4), 535-569.
- Ortega, D. L., Wang, H. H., Wu, L., & Hong, S. J. (2015). Retail channel and consumer demand for food quality in China. *China Economic Review*, 36, 359-366.
- Peel, D., & McLachlan, G. J. (2000). Robust mixture modelling using the t distribution. *Statistics and computing*, 10(4), 339-348.
- Parker, I. (2014). *Discourse Dynamics (Psychology Revivals): Critical Analysis for Social and Individual Psychology*. Routledge.
- Pickens, J. (2005). Attitudes and perceptions. *Organizational Behavior in Health Care*. Sudbury, MA: Jones and Bartlett Publishers, 43-75.
- Qu, Y., Tan, M., & Kutner, M. H. (1996). Random effects models in latent class analysis for evaluating accuracy of diagnostic tests. *Biometrics*, 797-810.
- Roy, R., Chintagunta, P. K., & Haldar, S. (1996). A framework for investigating habits, "The Hand of the Past," and heterogeneity in dynamic brand choice. *Marketing science*, 15(3), 280-299.
- Ryan, M., & Gerard, K. (2003). Using discrete choice experiments to value health care programs: current practice and future research reflections. *Applied health economics and health policy*, 2(1), 55-64.
- Scarpa, R., Campbell, D., & Hutchinson, W. G. (2007). Benefit estimates for landscape improvements: sequential Bayesian design and respondents' rationality in a choice experiment. *Land Economics*, 83(4), 617-634.

- Schwarz, G. (1978). Estimating the dimension of a model. *The annals of statistics*, 6(2), 461-464.
- Sclove, S. L. (1987). Application of model-selection criteria to some problems in multivariate analysis. *Psychometrika*, 52(3), 333-343.
- Steiger, J. H., Shapiro, A., & Browne, M. W. (1985). On the multivariate asymptotic distribution of sequential chi-square statistics. *Psychometrika*, 50(3), 253-263.
- Swait, J., & Adamowicz, W. (2001). The influence of task complexity on consumer choice: a latent class model of decision strategy switching. *Journal of Consumer Research*, 28(1), 135-148.
- Tanaka, M., Ida, T., Murakami, K., & Friedman, L. (2014). Consumers' willingness to pay for alternative fuel vehicles: A comparative discrete choice analysis between the US and Japan. *Transportation Research Part A: Policy and Practice*, 70, 194-209.
- Thoresen, T. O., & Vattø, T. E. (2015). Validation of the discrete choice labor supply model by methods of the new tax responsiveness literature. *Labour Economics*, 37, 38-53.
- Tofighi, D., & Enders, C. K. (2008). Identifying the correct number of classes in growth mixture models. *Advances in latent variable mixture models*, (Information Age Publishing, Inc.), 317-341.
- Vossler, C. A., Doyon, M., & Rondeau, D. (2012). Truth in consequentiality: theory and field evidence on discrete choice experiments. *American Economic Journal: Microeconomics*, 4(4), 145-171.
- Train, K. E. (1998). Recreation demand models with taste differences over people. *Land economics*, 230-239.
- Tversky, A. (1972). Elimination by aspects: A theory of choice. *Psychological review*, 79(4), 281.
- Vermunt, J. K., & Magidson, J. (2005). Latent GOLD 4.0 user's guide.
- Vermunt, J. K., & Magidson, J. (2013). Technical guide for Latent GOLD 5.0: Basic, advanced, and syntax. *Belmont, MA: Statistical Innovations Inc.*
- Whitty, J. A., Lancsar, E., Rixon, K., Golenko, X., & Ratcliffe, J. (2014). A systematic review of stated preference studies reporting public preferences for healthcare priority setting. *The Patient-Patient-Centered Outcomes Research*, 7(4), 365-386.
- Yang, C. C. (2006). Evaluating latent class analysis models in qualitative phenotype identification. *Computational Statistics & Data Analysis*, 50(4), 1090-1104.

Yang, C. C., & Yang, C. C. (2007). Separating latent classes by information criteria. *Journal of Classification*, 24(2), 183-203.

Yang, F. Y., & Tsai, C. C. (2008). Investigating university student preferences and beliefs about learning in the web-based context. *Computers & Education*, 50(4), 1284-1303.

Works Consulted

- Baetschmann, G., Staub, K. E., & Winkelmann, R. (2015). Consistent estimation of the fixed effects ordered logit model. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 178(3), 685-703.
- Bahamonde-Birke, F. J., Kunert, U., Link, H., & de Dios Ortúzar, J. (2015). About attitudes and perceptions: finding the proper way to consider latent variables in discrete choice models. *Transportation*, 1-19.
- Bozdogan, H. (1987). Model selection and Akaike's information criterion (AIC): The general theory and its analytical extensions. *Psychometrika*, 52(3), 345-370.
- Briesch, R. A., Chintagunta, P. K., & Matzkin, R. L. (2012). Nonparametric discrete choice models with unobserved heterogeneity. *Journal of Business & Economic Statistics*.
- Bulteel, K., Wilderjans, T. F., Tuerlinckx, F., & Ceulemans, E. (2013). CHull as an alternative to AIC and BIC in the context of mixtures of factor analyzers. *Behavior Research Methods*, 45(3), 782-791.
- Carson, R. T. (2000). Contingent valuation: a user's guide. *Environmental science & technology*, 34(8), 1413-1418.
- Coast, J., Al-Janabi, H., Sutton, E. J., Horrocks, S. A., Vosper, A. J., Swancutt, D. R., & Flynn, T. N. (2012). Using qualitative methods for attribute development for discrete choice experiments: issues and recommendations. *Health economics*, 21(6), 730-741.
- Cunha-e-Sá, M. A., Madureira, L., Nunes, L. C., & Otrachshenko, V. (2012). Protesting and justifying: a latent class model for contingent valuation with attitudinal data. *Environmental and Resource Economics*, 52(4), 531-548.
- de Bekker-Grob, E. (2009). Discrete choice experiments in health care: theory and applications.
- de Oña, J., López, G., Mujalli, R., & Calvo, F. J. (2013). Analysis of traffic accidents on rural highways using Latent Class Clustering and Bayesian Networks. *Accident Analysis & Prevention*, 51, 1-10.
- Dubé, J. P., Hitsch, G. J., & Rossi, P. E. (2010). State dependence and alternative explanations for consumer inertia. *The RAND Journal of Economics*, 41(3), 417-445.
- Field, A. (2013). *Discovering statistics using IBM SPSS statistics*. Sage.
- Gelman, A., Hwang, J., & Vehtari, A. (2014). Understanding predictive information criteria for Bayesian models. *Statistics and Computing*, 24(6), 997-1016.

Greene, W. H., & Hensher, D. A. (2013). Revealing additional dimensions of preference heterogeneity in a latent class mixed multinomial logit model. *Applied Economics*, 45(14), 1897-1902.

Hanemann, W. M. (1994). Valuing the environment through contingent valuation. *The Journal of Economic Perspectives*, 8(4), 19-43.

Haan, P., Kemptner, D., & Uhlendorff, A. (2015). Bayesian procedures as a numerical tool for the estimation of an intertemporal discrete choice model. *Empirical Economics*, 49(3), 1123-1141.

Hausman, J. (2012). Contingent valuation: from dubious to hopeless. *The Journal of Economic Perspectives*, 26(4), 43-56.

Hess, S., Stathopoulos, A., & Daly, A. (2012). Allowing for heterogeneous decision rules in discrete choice models: an approach and four case studies. *Transportation*, 39(3), 565-591.

Hoehn, J. P., & Randall, A. (1987). A satisfactory benefit cost indicator from contingent valuation. *Journal of environmental Economics and Management*, 14(3), 226-247.

Jonker, M., & Van der Vaart, A. (2014). On the correction of the asymptotic distribution of the likelihood ratio statistic if nuisance parameters are estimated based on an external source. *The international journal of biostatistics*, 10(2), 123-142.

Keller, B. A., & Hirsch, C. R. (1998). Student preferences for representations of functions. *International Journal of Mathematical Education in Science and Technology*, 29(1), 1-17.

Koning, R. H., & Ridder, G. (2003). Discrete choice and stochastic utility maximization. *The Econometrics Journal*, 6(1), 1-27.

Lagarde, M. (2013). Investigating Attribute Non-Attendance and Its Consequences in Choice Experiments with Latent Class Models. *Health economics*, 22(5), 554-567.

Li, W., Nachtsheim, C. J., Wang, K., Reul, R., & Albrecht, M. (2013). Conjoint analysis and discrete choice experiments for quality improvement. *Journal of Quality Technology*, 45(1), 74.

Lin, T. H., & Dayton, C. M. (1997). Model selection information criteria for non-nested latent class models. *Journal of Educational and Behavioral Statistics*, 22(3), 249-264.

Loomis, J., Kent, P., Strange, L., Fausch, K., & Covich, A. (2000). Measuring the total

economic value of restoring ecosystem services in an impaired river basin: results from a contingent valuation survey. *Ecological economics*, 33(1), 103-117.

Lo, Y., Mendell, N. R., & Rubin, D. B. (2001). Testing the number of components in a normal mixture. *Biometrika*, 88(3), 767-778.

Lovreglio, R., Borri, D., dell'Olio, L., & Ibeas, A. (2014). A discrete choice model based on random utilities for exit choice in emergency evacuations. *Safety science*, 62, 418-426.

Mahieu, P. A., Andersson, H., Beaumais, O., Crastes, R., & Wolff, F. C. (2014). Is Choice Experiment Becoming More Popular than Contingent Valuation? A Systematic Review in Agriculture, Environment and Health (No. 2014.12).

McFadden, D. (1997). *Why is natural resource damage assessment so hard?* Department of Agricultural and Applied Economics, College of Agricultural and Life Sciences, University of Wisconsin--Madison.

McFadden, D. (2000). Disaggregate behavioral travel demand's RUM side. *Travel Behaviour Research*, 17-63.

McFadden, D., & Train, K. (2000). Mixed MNL models for discrete response. *Journal of applied Econometrics*, 15(5), 447-470.

Moon, H. R., & Schorfheide, F. (2012). Bayesian and frequentist inference in partially identified models. *Econometrica*, 80(2), 755-782.

Morey, E. R., Breffle, W. S., & Greene, P. A. (2001). Two Nested Constant-Elasticity-of-Substitution Models of Recreational Participation and Site Choice: An 'Alternatives' Model and an 'Expenditures' Model. *American Journal of Agricultural Economics*, 83(2), 414-427.

Morey, E. R., Breffle, W. S., Rowe, R. D., & Waldman, D. M. (2002). Estimating recreational trout fishing damages in Montana's Clark Fork River basin: summary of a natural resource damage assessment. *Journal of Environmental Management*, 66(2), 159-170.

Nagin, D. S. (1999). Analyzing developmental trajectories: a semiparametric, group-based approach. *Psychological methods*, 4(2), 139.

Norets, A., & Tang, X. (2013). Semiparametric Inference in dynamic binary choice models. *The Review of Economic Studies*, rdt050.

Ortuzar, J. D. D., & Willumsen, L. G. (2002). *Modelling transport*. John Wiley & Sons. West Sussex, England.

- Rissanen, J. (1978). Modeling by shortest data description. *Automatica*, 14(5), 465-471.
- Stigka, E. K., Paravantis, J. A., & Mihalakakou, G. K. (2014). Social acceptance of renewable energy sources: A review of contingent valuation applications. *Renewable and Sustainable Energy Reviews*, 32, 100-106.
- Sung, M. C., McDonald, D. C., & Johnson, J. E. (2016). Probabilistic forecasting with discrete choice models: Evaluating predictions with pseudo-coefficients of determination. *European Journal of Operational Research*, 248(3), 1021-1030.
- Vermunt, J. K. (2010). Latent class modeling with covariates: Two improved three-step approaches. *Political analysis*, 18(4), 450-469.
- Wasserman, L. (2000). Bayesian model selection and model averaging. *Journal of mathematical psychology*, 44(1), 92-107.
- Wedel, M., DeSarbo, W. S., Bult, J. R., & Ramaswamy, V. (1993). A latent class Poisson regression model for heterogeneous count data. *Journal of Applied Econometrics*, 8(4), 397-411.
- Wilcox, N. (2015). Error and generalization in discrete choice under risk (No. 15-11).
- Wu, Q. (2009). *Class extraction and classification accuracy in latent class models* (Doctoral dissertation, The Pennsylvania State University).
- Yelland, L. N., Salter, A. B., & Ryan, P. (2011). Performance of the modified Poisson regression approach for estimating relative risks from clustered prospective data. *American journal of epidemiology*, kwr183.

Appendices

Appendix A: MTU Enrollment Numbers

The following table shows the MTU enrollment numbers from the years 2009-2016.

Table A1: Enrollment in Natural Resource Courses (2009-2016)

Undergrad									Total
	2009	2010	2011	2012	2013	2014	2015	2016	Undergrad
Natural resource	7	4	5	7	9	14	14	-	60
Environmental	-	14	12	25	11	13	7	12	94
Energy	-	23	19	32	28	32	18	21	173
Mineral industry	-	-	7	11	7	4	16	13	58
									Total
Total		44	42	73	53	58	55	60	385
Graduate									
	2009	2010	2011	2012	2013	2014	2015	2016	Graduate
Natural resource	15	15	12	9	6	6	9	-	72
Environmental	-	9	7	3	7	1	4	4	35
Energy	-	19	13	19	25	8	7	3	94
Mineral industry	-	-	7	2	5	3	2	2	21
									Total
Total		43	42	36	46	18	19	18	222

Appendix B: Investigating Random Choice Responses

Both economic and psychological research studies have been done in an attempt to better understand the human decision making process. The first noteworthy study asserts a few different indicators for the researchers to consider when examining the decision making process of respondents. Bahmonde-Birke et al. (2015) broke this down into two simple components in their latent class experiment; attitudinal-based and perception-based decision making. The former is a function of a subjects' historical experiences as well as their temperament, creating a "mind-set" formed from the knowledge developed from those experiences (Pickens, 2005). Conversely, perception is how a subject experiences their environment and thus is a function of both the subject and the incitement of their environment. Translating the latter, one can say the attitudes are constant across all alternatives, and perceptual latent variables should be dynamic across each subject and alternatives.

When individuals are forced to choose among several alternatives, they often display some sort of uncertainty which can cause people to make different or inconsistent choices even when faced with the same alternatives a second time (Tversky, 1972). Louviere (2006) asserts how the individuals under observation could be detrimental to a CE if they are placed in an arbitrary setting (e.g., car auctions). This is because the subjects are placed into an unusual environment that may cause them to act irrationally; such action can often lead to unreliably biased responses. Using LCA allows for class specification via heterogeneity and therefore could consist of a class where respondents are essentially answering randomly or aimlessly. Morey and Thacher (2012) demonstrate in their latent-class study that there are discrepancies in individuals' interpretation of hypothetical questioning. They propose some individuals (especially economists) rationalize not "wasting" time answering challenging hypothetical questions and others consider hypothetical questioning "only" worthy of hypothetical answering; implying they are mindlessly and effortlessly answering questions (Morey and Thacher, 2012).

Many psychological studies of individual decision making suggest that while the decision is made by that individual there are hidden factors that also influence one's choice. People make decisions but not fully at their own choosing (Tversky, 1972; Parker, 2014).

This brings a whole world of latent variables that could be influencing ones' decision to simply choose A over B. For example, a simple article on the impact of global warming could cause an individual to choose a career in environmental economics. Another, more simple example would be a students' parental political preference influencing that individual to choose a conservative preference, when their actual beliefs support a more liberal party. A recent publication by Parker (2014) in a psychiatric analysis describes a dynamic between an individual's choice and companies', authorities' and parents' influences on what is taught in academia. This suggests that an individual's choice is never 100 percent their own, as external factors, whether weak or strong, are playing some roll in the ultimate choice.

The variability of influences is numerous, however available information can be used to design a set of important alternatives and social factors that can exploit the most important attributes and additionally disclose some choice heterogeneity. All the interpersonal variables that drive a person's choice are unobservable, but regardless the reason for the choice; DCMs can be constructed based on the limited information available by a SP instrument.

Appendix C: Survey Versions 1-3

Surveys versions are presented below in ascending order.

Ten-minute Questionnaire about Economics Courses on the Natural Environment at Michigan Tech

Q1 – Michigan Tech’s School of Business and Economics offers **four (4) economics courses** focusing on the natural environment (e.g., *environmental economics*, *natural resource economics*, *energy economics*, *mineral industry economics*). Prior to this survey, how aware were you that these electives are available?

1 – Not at all aware 2 – A little aware 3 – Somewhat aware 4 – Very aware 5 – Extremely aware

Q2 – Michigan Tech’s School of Business and Economics also offers a **Master’s Degree in Applied Natural Resource Economics**. Classes such as those above are taught at the graduate level, and other support electives such as mathematical economics, econometrics, and advanced microeconomics are offered as well. Prior to this survey, how aware were you about this master’s program?

1 – Not at all aware 2 -- A little aware 3 – Somewhat aware 4 – Very aware 5 – Extremely aware

Q3 – What is the likelihood you would enroll in one of the **four (4) courses** in Q1, or **the master’s program**? (Choose 5 if you already have taken one of these courses, or are currently enrolled in one of these courses or the master’s program)

1 – Not at all likely 2 -- A little likely 3 – Somewhat likely 4 – Very likely 5 – Extremely likely

Michigan Tech’s School of Business and Economics is interested in learning more about the attributes of **the four (4) economics courses** focusing on the natural environment that affect enrollment. The next six questions (Q4 through Q9) each present two (2) alternative courses with varying attributes. Consider both courses, and at the bottom of each choice pick the course you would prefer (even if you do not like either of them, or like both of them, pick either course A or B). Assume all of the attributes are given; you cannot change them. Some comparisons might be easy, while others might be quite difficult.

Q4 Please pick the alternative you prefer (course A or course B):

Attributes

Focus	Conservation/sustainability	Resource use/energy markets
Credit	Graduation only	Graduation only
Study hours	7 hours	3 hours
Teacher	Tenured professor	Instructor/lecturer
Paper	Small portion of grade	Small portion of grade
Time	Late afternoon	Late afternoon
Day	MWF	Tu/Th
Semester	Spring	Spring

I prefer option: (circle one)

A

B

Q5 Please pick the alternative you prefer (course A or course B):

Attributes

Focus	Conservation/sustainability	Resource use/energy markets
Credit	Major requirement fulfilled	Graduation only
Study hours	5 hours	7 hours
Teacher	Instructor/lecturer	Untenured junior professor
Paper	Large portion of grade	Small portion of grade
Time	Early morning	Early morning
Day	Tu/Th	Tu/Th
Semester	Fall	Fall

I prefer option: (circle one)

A

B

Q6 Please pick the alternative you prefer (course A or course B):

Attributes

Focus	Balance use and conservation	Conservation/sustainability
Credit	Graduation only	Graduation only
Study hours	3 hours	5 hours
Teacher	Tenured professor	Instructor/lecturer
Paper	Large portion of grade	Small portion of grade
Time	Early morning	Early morning
Day	Tu/Th	MWF
Semester	Spring	Fall

I prefer option: (circle one)

A

B

Q7 Please pick the alternative you prefer (course A or course B):

Attributes

Focus	Balance use and conservation	Balance use and conservation
Credit	Graduation only	Major requirement fulfilled
Study hours	7 hours	3 hours
Teacher	Instructor/lecturer	Tenured professor
Paper	Large portion of grade	Small portion of grade
Time	Midday	Early morning
Day	MWF	MWF
Semester	Fall	Fall

I prefer option: (circle one)

A

B

Q8 Please pick the alternative you prefer (course A or course B):

Attributes

Focus	Balance use and conservation	Balance use and conservation
Credit	Graduation only	Major requirement fulfilled
Study hours	5 hours	7 hours
Teacher	Untenured junior professor	Instructor/lecturer
Paper	Large portion of grade	Small portion of grade
Time	Late afternoon	Midday
Day	MWF	Tu/Th
Semester	Spring	Fall

I prefer option: (circle one)

A

B

Q9 Please pick the alternative you prefer (course A or course B):

Attributes

Focus	Conservation/sustainability	Balance use and conservation
Credit	Graduation only	Graduation only
Study hours	5 hours	7 hours
Teacher	Instructor/lecturer	Instructor/lecturer
Paper	Small portion of grade	Large portion of grade
Time	Early morning	Midday
Day	MWF	MWF
Semester	Spring	Spring

I prefer option: (circle one)

A

B

Michigan Tech's School of Business and Economics is interested in learning more about the types of students who may be interested in taking these courses. The next questions are about you, the student:

Q10: Is your status [freshman-or-sophomore], [junior-or-senior], or [graduate student]? (circle one)

Q11: Are you a male or a female? (circle one)

Q12: Are you a domestic student or foreign exchange student? (circle one)

Q13: Is your major HASS (humanities, arts, or soc. sciences) or STEM (science, tech, eng., or math) or other? (circle one)

Q14: Is your interest in economics low, medium, or high? (circle one)

Q15: After graduation, are you more likely to go to graduate school or into the job market or other? (circle one)

Q16: When you do go on the job market, will you likely seek a job in business or government or other? (circle one)

Q17: Politically, do you consider yourself to be more liberal or more conservative? (circle one)

ID (student leave blank) _____

Thank you for participating in our survey!

Version 1

Ten-minute Questionnaire about Economics Courses on the Natural Environment at Michigan Tech

Q1 – Michigan Tech’s School of Business and Economics offers **four (4) economics courses** focusing on the natural environment (e.g., *environmental economics*, *natural resource economics*, *energy economics*, *mineral industry economics*). Prior to this survey, how aware were you that these electives are available?

1 – Not at all aware 2 – A little aware 3 – Somewhat aware 4 – Very aware 5 – Extremely aware

Q2 – Michigan Tech’s School of Business and Economics also offers a **Master’s Degree in Applied Natural Resource Economics**. Classes such as those above are taught at the graduate level, and other support electives such as mathematical economics, econometrics, and advanced microeconomics are offered as well. Prior to this survey, how aware were you about this master’s program?

1 – Not at all aware 2 -- A little aware 3 – Somewhat aware 4 – Very aware 5 – Extremely aware

Q3 – What is the likelihood you would enroll in one of the **four (4) courses** in Q1, or **the master's program**? (Choose 5 if you already have taken one of these courses, or are currently enrolled in one of these courses or the master's program)

1 – Not at all likely 2 -- A little likely 3 – Somewhat likely 4 – Very likely 5 – Extremely likely

Michigan Tech's School of Business and Economics is interested in learning more about the attributes of **the four (4) economics courses** focusing on the natural environment that affect enrollment. The next six questions (Q4 through Q9) each present two (2) alternative courses with varying attributes. Consider both courses, and at the bottom of each choice pick the course you would prefer (even if you do not like either of them, or like both of them, pick either course A or B). Assume all of the attributes are given; you cannot change them. Some comparisons might be easy, while others might be quite difficult.

Q4 Please pick the alternative you prefer (course A or course B):

Attributes		
Focus	Resource use/energy markets	Resource use/energy markets
Credit	Graduation only	Major requirement fulfilled
Study hours	5 hours	7 hours
Teacher	Tenured professor	Untenured junior professor
Paper	Large portion of grade	Large portion of grade
Time	Midday	Early morning
Day	MWF	Tu/Th
Semester	Spring	Spring

I prefer option: (circle one)

A

B

Q5 Please pick the alternative you prefer (course A or course B):

Attributes		
Focus	Conservation/sustainability	Resource use/energy markets
Credit	Major requirement fulfilled	Major requirement fulfilled
Study hours	7 hours	5 hours
Teacher	Tenured professor	Tenured professor
Paper	Large portion of grade	Small portion of grade
Time	Late afternoon	Midday
Day	Tu/Th	MWF
Semester	Fall	Spring

I prefer option: (circle one)

A

B

Q6 Please pick the alternative you prefer (course A or course B):

Attributes		
Focus	Balance use and conservation	Resource use/energy markets
Credit	Graduation only	Major requirement fulfilled
Study hours	3 hours	3 hours
Teacher	Tenured professor	Instructor/lecturer
Paper	Large portion of grade	Large portion of grade
Time	Early morning	Late afternoon
Day	Tu/Th	MWF
Semester	Spring	Fall

I prefer option: (circle one)

A

B

Q7 Please pick the alternative you prefer (course A or course B):

Attributes

Focus	Conservation/sustainability	Conservation/sustainability
Credit	Major requirement fulfilled	Major requirement fulfilled
Study hours	3 hours	5 hours
Teacher	Untenured junior professor	Instructor/lecturer
Paper	Large portion of grade	Large portion of grade
Time	Midday	Early morning
Day	MWF	Tu/Th
Semester	Fall	Spring

I prefer option: (circle one)

A

B

Q8 Please pick the alternative you prefer (course A or course B):

Attributes

Focus	Balance use and conservation	Balance use and conservation
Credit	Major requirement fulfilled	Major requirement fulfilled
Study hours	5 hours	5 hours
Teacher	Untenured junior professor	Untenured junior professor
Paper	Small portion of grade	Large portion of grade
Time	Late afternoon	Late afternoon
Day	Tu/Th	MWF
Semester	Spring	Fall

I prefer option: (circle one)

A

B

Q9 Please pick the alternative you prefer (course A or course B):

Attributes

Focus	Resource use/energy markets	Conservation/sustainability
Credit	Major requirement fulfilled	Graduation only
Study hours	5 hours	3 hours
Teacher	Tenured professor	Untenured junior professor
Paper	Small portion of grade	Small portion of grade
Time	Midday	Midday
Day	Tu/Th	Tu/Th
Semester	Fall	Spring

I prefer option: (circle one)

A

B

Michigan Tech's School of Business and Economics is interested in learning more about the types of students who may be interested in taking these courses. The next questions are about you, the student:

Q10: Is your status [freshman-or-sophomore], [junior-or-senior], or [graduate student]? (circle one)

Q11: Are you a male or a female? (circle one)

Q12: Are you a domestic student or foreign exchange student? (circle one)

Q13: Is your major HASS (humanities, arts, or soc. sciences) or STEM (science, tech, eng., or math) or other? (circle one)

Q14: Is your interest in economics low, medium, or high? (circle one)

Q15: After graduation, are you more likely to go to graduate school or into the job market or other? (circle one)

Q16: When you do go on the job market, will you likely seek a job in business or government or other? (circle one)

Q17: Politically, do you consider yourself to be more liberal or more conservative? (circle one)

ID (student leave blank) _____

Thank you for participating in our survey!

Version 2

Ten-minute Questionnaire about Economics Courses on the Natural Environment at Michigan Tech

Q1 – Michigan Tech’s School of Business and Economics offers **four (4) economics courses** focusing on the natural environment (e.g., *environmental economics*, *natural resource economics*, *energy economics*, *mineral industry economics*). Prior to this survey, how aware were you that these electives are available?

1 – Not at all aware 2 – A little aware 3 – Somewhat aware 4 – Very aware 5 – Extremely aware

Q2 – Michigan Tech’s School of Business and Economics also offers a **Master’s Degree in Applied Natural Resource Economics**. Classes such as those above are taught at the graduate level, and other support electives such as mathematical economics, econometrics, and advanced microeconomics are offered as well. Prior to this survey, how aware were you about this master’s program?

1 – Not at all aware 2 -- A little aware 3 – Somewhat aware 4 – Very aware 5 – Extremely aware

Q3 – What is the likelihood you would enroll in one of the **four (4) courses** in Q1, or **the master’s program**? (Choose 5 if you already have taken one of these courses, or are currently enrolled in one of these courses or the master’s program)

1 – Not at all likely 2 -- A little likely 3 – Somewhat likely 4 – Very likely 5 – Extremely likely

Michigan Tech’s School of Business and Economics is interested in learning more about the attributes of **the four (4) economics courses** focusing on the natural environment that affect enrollment. The next six questions (Q4 through Q9) each present two (2) alternative courses with varying attributes. Consider both courses, and at the bottom of each choice pick the course you would prefer (even if you do not like either of them, or like both of them, pick either course A or B). Assume all of the attributes are given; you cannot change them. Some comparisons might be easy, while others might be quite difficult.

Q4 Please pick the alternative you prefer (course A or course B):

Attributes

Focus	Resource use/energy markets	Resource use/energy markets
Credit	Major requirement fulfilled	Graduation only
Study hours	7 hours	3 hours
Teacher	Untenured junior professor	Instructor/lecturer
Paper	Small portion of grade	Small portion of grade
Time	Early morning	Late afternoon
Day	MWF	Tu/Th
Semester	Spring	Fall

I prefer option: (circle one)

A

B

Q5 Please pick the alternative you prefer (course A or course B):

Attributes

Focus	Balance use and conservation	Balance use and conservation
Credit	Graduation only	Major requirement fulfilled
Study hours	5 hours	7 hours
Teacher	Untenured junior professor	Instructor/lecturer
Paper	Small portion of grade	Small portion of grade
Time	Late afternoon	Midday
Day	Tu/Th	Tu/Th
Semester	Fall	Spring

I prefer option: (circle one)

A

B

Q6 Please pick the alternative you prefer (course A or course B):

Attributes

Focus	Conservation/sustainability	Resource use/energy markets
Credit	Major requirement fulfilled	Graduation only
Study hours	7 hours	5 hours
Teacher	Tenured professor	Tenured professor
Paper	Large portion of grade	Large portion of grade
Time	Late afternoon	Midday
Day	Tu/Th	Tu/Th
Semester	Spring	Fall

I prefer option: (circle one)

A

B

Q7 Please pick the alternative you prefer (course A or course B):

Attributes

Focus	Balance use and conservation	Resource use/energy markets
Credit	Major requirement fulfilled	Major requirement fulfilled
Study hours	3 hours	3 hours
Teacher	Tenured professor	Instructor/lecturer
Paper	Small portion of grade	Large portion of grade
Time	Early morning	Late afternoon
Day	MWF	MWF
Semester	Fall	Spring

I prefer option: (circle one)

A

B

Q8 Please pick the alternative you prefer (course A or course B):

Attributes

Focus	Conservation/sustainability	Resource use/energy markets
Credit	Major requirement fulfilled	Graduation only
Study hours	3 hours	7 hours
Teacher	Untenured junior professor	Untenured junior professor
Paper	Small portion of grade	Large portion of grade
Time	Midday	Early morning
Day	MWF	MWF
Semester	Spring	Fall

I prefer option: (circle one)

A

B

Q9 Please pick the alternative you prefer (course A or course B):

Attributes

Focus	Conservation/sustainability	Conservation/sustainability
Credit	Graduation only	Graduation only
Study hours	7 hours	3 hours
Teacher	Tenured professor	Untenured junior professor
Paper	Small portion of grade	Large portion of grade
Time	Late afternoon	Midday
Day	MWF	Tu/Th
Semester	Fall	Fall

I prefer option: (circle one)

A

B

Michigan Tech's School of Business and Economics is interested in learning more about the types of students who may be interested in taking these courses. The next questions are about you, the student:

Q10: Is your status [freshman-or-sophomore], [junior-or-senior], or [graduate student]? (circle one)

Q11: Are you a male or a female? (circle one)

Q12: Are you a domestic student or foreign exchange student? (circle one)

Q13: Is your major HASS (humanities, arts, or soc. sciences) or STEM (science, tech, eng., or math) or other? (circle one)

Q14: Is your interest in economics low, medium, or high? (circle one)

Q15: After graduation, are you more likely to go to graduate school or into the job market or other? (circle one)

Q16: When you do go on the job market, will you likely seek a job in business or government or other? (circle one)

Q17: Politically, do you consider yourself to be more liberal or more conservative? (circle one)

ID (student leave blank) _____

Thank you for participating in our survey!

Version 3

Appendix D: Latent Gold Choice

Choice Format

As previously mentioned, there are no “opt-out” alternatives so each respondent has to (1) make a choice between the binary choice options provided, or (2) not answer the question at all. Regardless the reason a respondent failed to respond, all unanswered questions are recorded in the data set as -88. This number makes for easily identifying unanswered questions even though there is no significant effect of these missing variables on the model. -88 is additionally used to denote unanswered socio-demographic questions. These socio-demographic questions are designed to give us information on the respondent’s individual characteristics such as: sex, academic rank, nationality, etc. These questions are uniform throughout the three survey versions.

Unconditional Probabilities

Denoted as “Profile” in LGC are the unconditional probabilities of attributes. These probabilities are indicators of the strength of the relationship between each variable and each class. The sum of the probabilities for each variable are equal to 1 in the respective class. For example, in Table A2 below, conditional on being in Class 4, there is a 99% chance that a respondent from class 4 will choose to have class on T/TH and a 1% chance they will choose MWF course. The latter example is bolded in the subsequent table (see Table A2).

Table A2: Unconditional Probabilities

	Class1	Class2	Class3	Class4
Class Size	0.399	0.2915	0.1853	0.1242
Attributes				
<i>focus</i>				
RESOURCE USE/ENERGY MARKETS	0.1049	0.5468	0.0343	0.0064
CONSERVATION/ SUSTAINABILITY	0.0745	0.2888	0.1266	0.9936
BALANCE USE AND CONSERVATION	0.8206	0.1643	0.839	0
<i>major</i>				
GRADUATION ONLY	0.2148	0.4053	0.2755	0.9521
MAJOR REQUIREMENT FULFILLED	0.7852	0.5947	0.7245	0.0479
<i>studyhr</i>				
HIGH (7HRS)	0.023	0.2957	0.3906	0.1268
MED (5HRS)	0.8601	0.3499	0.2918	0.8707
LOW (3HRS)	0.1169	0.3544	0.3176	0.0025
<i>profrank</i>				
INSRUCTOR/LECTURER	0.335	0.2874	0.6495	0.0016
JUNIOR PROFESSOR	0.1805	0.3407	0.3217	0.0001
TENURED PROFESSOR	0.4845	0.372	0.0288	0.9983
<i>smpaper</i>				
LARGE % OF GRADE	0.089	0.5127	0.0354	0.4095
SMALL % OF GRADE	0.911	0.4873	0.9646	0.5905
<i>time</i>				
MORNING	0.3399	0.1357	0.7934	0.0083
MIDDAY	0.3389	0.4111	0.1879	0.916
LATE AFTERNOON	0.3212	0.4532	0.0187	0.0756
<i>MWF</i>				
T/TH	0.167	0.5723	0.714	0.99
MWF	0.833	0.4277	0.286	0.01
<i>fall</i>				
SPRING	0.6478	0.5711	0.0415	0.0256
FALL	0.3522	0.4289	0.9585	0.9744

Table A2 continues on the next page.

Table A2 Continued: Unconditional Probabilities

Covariates				
	Class1	Class2	Class3	Class4
<i>INTEREST</i>				
LOW	0.2438	0.1171	0.1667	0.3517
MED	0.6216	0.621	0.7011	0.5094
HIGH	0.1346	0.2619	0.1322	0.1389
<i>POLPREF</i>				
CON	0.6819	0.6119	0.6317	0.2132
LIB	0.3181	0.3881	0.3683	0.7868
<i>USA</i>				
FOREIGN	0.0197	0.1128	0.0007	0.0046
DOMESTIC	0.9803	0.8872	0.9993	0.9954

Appendix E: LGC Files: Design Matrix and Choice Sets

Table A3 displays this study's original design matrix. This orthogonal design was created by Jennifer A. Thacher; graduate director and associate professor at the of University New Mexico. She is an applied micro-economist with a specialization in environmental economics. Further, Thacher has much previous work with heterogeneous models, including LCA. Some of her publications include:

Aldrich, G. A., Grimsrud, K. M., Thacher, J. A., & Kotchen, M. J. (2007). Relating environmental attitudes and contingent values: how robust are methods for identifying preference heterogeneity? *Environmental and Resource Economics*, 37(4), 757-775.

Hand, M. S., Thacher, J. A., McCollum, D. W., & Berrens, R. P. (2008). Intra-regional amenities, wages, and home prices: the role of forests in the Southwest. *Land Economics*, 84(4), 635-651.

Morey, E., Thacher, J., & Breffle, W. (2006). Using angler characteristics and attitudinal data to identify environmental preference classes: a latent-class model. *Environmental and Resource Economics*, 34(1), 91-115.

Morey, E., Thacher, J. A., & Craighead, W. E. (2007). Patient preferences for depression treatment programs and willingness to pay for treatment. *Journal of Mental Health Policy and Economics*, 10(2), 73.

Thacher, J. A., Morey, E., & Craighead, W. E. (2005). Using patient characteristics and attitudinal data to identify depression treatment preference groups: a latent-class model. *Depression and anxiety*, 21(2), 47-54.

Table A3: Original Design Matrix

choice code	cons.	bal.	maj	hwlo	hwmid	junpro	ten	smpap	mid	late	MWF	fall
cc1	1	0	0	0	0	0	1	1	0	1	1	0
cc2	0	0	0	1	0	0	0	1	0	1	0	0
cc3	1	0	1	0	1	0	0	0	0	0	0	1
cc4	0	0	0	0	0	1	0	1	0	0	0	1
cc5	0	1	0	1	0	0	1	0	0	0	0	0
cc6	1	0	0	0	1	0	0	1	0	0	1	1
cc7	0	1	0	0	0	0	0	0	1	0	1	1
cc8	0	1	1	1	0	0	1	1	0	0	1	1
cc9	0	1	0	0	1	1	0	0	0	1	1	0
cc10	0	1	1	0	0	0	0	1	1	0	0	1
cc11	1	0	0	0	1	0	0	1	0	0	1	0
cc12	0	1	0	0	0	0	0	0	1	0	1	0
cc13	0	0	0	0	1	0	1	0	1	0	1	0
cc14	0	0	1	0	0	1	0	0	0	0	0	0
cc15	1	0	1	0	0	0	1	0	0	1	0	1
cc16	0	0	1	0	1	0	1	1	1	0	1	0
cc17	0	1	0	1	0	0	1	0	0	0	0	0
cc18	0	0	1	1	0	0	0	0	0	1	1	1
cc19	1	0	1	1	0	1	0	0	1	0	1	1
cc20	1	0	1	1	0	0	0	0	0	0	0	0
cc21	0	1	1	0	1	1	0	1	0	1	0	0
cc22	0	1	1	0	1	1	0	0	0	1	1	1
cc23	0	0	1	0	1	0	1	1	1	0	0	1
cc24	1	0	0	1	0	1	0	1	1	0	0	0
cc25	0	0	1	0	0	1	0	1	0	0	1	0
cc26	0	0	0	1	0	0	0	1	0	1	0	1
cc27	0	1	0	0	1	1	0	1	0	1	0	1
cc28	0	1	1	0	0	0	0	1	1	0	0	0
cc29	1	0	1	0	0	0	1	0	0	1	0	0
cc30	0	0	0	0	1	0	1	0	1	0	0	1
cc31	0	1	1	1	0	0	1	1	0	0	1	1
cc32	0	0	1	1	0	0	0	0	0	1	1	0
cc33	1	0	1	1	0	1	0	1	1	0	1	0
cc34	0	0	0	0	0	1	0	0	0	0	1	1
cc35	1	0	0	0	0	0	1	1	0	1	1	1
cc36	1	0	0	1	0	1	0	0	1	0	0	1

Choice Sets

Table A4 illustrates the alternative choice set file used for estimation in LGC. There are 36 alternatives (see Table A3). In binomial choice sets, where respondents choose 0 or 1, that results in 18 different choice sets that correspond to the alternatives file; based on the alternative chosen. These choice sets are shown below in Table A4, and is one of the three required files by LGC to estimate a LCM such as the one in this study. The second is discussed subsequently, and the third is the massive choice response data not presented.

Table A4: Choice Sets

Set ID	alt1	alt2
1	cc1	cc2
2	cc3	cc4
3	cc5	cc6
4	cc7	cc8
5	cc9	cc10
6	cc11	cc12
7	cc13	cc14
8	cc15	cc16
9	cc17	cc18
10	cc19	cc20
11	cc21	cc22
12	cc23	cc24
13	cc25	cc26
14	cc27	cc28
15	cc29	cc30
16	cc31	cc32
17	cc33	cc34
18	cc35	cc36

Alternative File

Our alternatives file used in LGC consists of a discrete 0, 1, 2 scaled matrix representing each possible alternative of course attributes, varying only in attribute levels. Table A5 below represents an alternative, recoded, design matrix of the one shown in Table A3. The matrix in Table A5 incorporates attribute levels into the respective variable, rather than have them all coded as binomial variables as in Table A3. This was a requirement for

LGC in order to estimate a LCM. This was used as the “alternatives” file for LGC. It is number 2 of 3 files. A key follows Table A5 for interpretation of attribute levels.

Table A5: LGC Alternatives Matrix

choice code	focus	major	studyhr	rank	smpaper	time	MWF	fall
cc1	1	0	0	2	1	2	1	0
cc2	0	0	1	0	1	2	0	0
cc3	1	1	2	0	0	0	0	1
cc4	0	0	0	1	1	0	0	1
cc5	2	0	1	2	0	0	0	0
cc6	1	0	2	0	1	0	1	1
cc7	2	0	0	0	0	1	1	1
cc8	2	1	1	2	1	0	1	1
cc9	2	0	2	1	0	2	1	0
cc10	2	1	0	0	1	1	0	1
cc11	1	0	2	0	1	0	1	0
cc12	2	0	0	0	0	1	1	0
cc13	0	0	2	2	0	1	1	0
cc14	0	1	0	1	0	0	0	0
cc15	1	1	0	2	0	2	0	1
cc16	0	1	2	2	1	1	1	0
cc17	2	0	1	2	0	0	0	0
cc18	0	1	1	0	0	2	1	1
cc19	1	1	1	1	0	1	1	1
cc20	1	1	1	0	0	0	0	0
cc21	2	1	2	1	1	2	0	0
cc22	2	1	2	1	0	2	1	1
cc23	0	1	2	2	1	1	0	1
cc24	1	0	1	1	1	1	0	0
cc25	0	1	0	1	1	0	1	0
cc26	0	0	1	0	1	2	0	1
cc27	2	0	2	1	1	2	0	1
cc28	2	1	0	0	1	1	0	0
cc29	1	1	0	2	0	2	0	0
cc30	0	0	2	2	0	1	0	1
cc31	2	1	1	2	1	0	1	1
cc32	0	1	1	0	0	2	1	0
cc33	1	1	1	1	1	1	1	0
cc34	0	0	0	1	0	0	1	1
cc35	1	0	0	2	1	2	1	1
cc36	1	0	1	1	0	1	0	1

Table A5 Continued: LGC Alternatives Matrix

Coding	rank	focus	studyhr	time
0	instructor	resource	hi (7hr)	early
1	junprof	conserve	mid (5hr)	mid
2	tenure	balance	lo (3hr)	late

Appendix F: Student Awareness Correlations

Table A6 investigates the relationship between awareness and likelihood to enroll by using a correlation matrix to analyze the first three survey questions (see Section 4.4). Table A7 displays a significant relationship between student likelihood of enrollment and their level of economic interest. Correlations that are significant at the 0.01 level($\alpha = .01$) are bolded and further denoted in the table.

The Spearman rho correlation was used in this experiment because of the non-parametric nature of the data. It is typical to see Pearson's correlations however in the case of using ranking scales it is best to measure the monotonic relationship of variables, as Spearman coefficients do.¹⁵ Pearson correlations measure the linear relationship between variables and are used for data on an interval scale. It additionally holds the assumption that the data is normally distributed, which is not the case in this study.

¹⁵ Monotonic in this sense means go up in one direction only or stays at the same level.

Table A6: Correlation Matrix of Questions 1-3

Awareness and Likelihood of Enrollment Correlations				
Spearman's rho		Awareness of NR Economics Courses	Awareness of BS NRE Master's Program	Likelihood of Enrollment in Q1 or Q2
Awareness of NR Economics Courses	Correlation Coefficient	1.000	.476**	.368**
	Sig. (2-tailed)		0.000	0.000
	N	4338	4338	4332
Awareness of BS NRE Master's Program	Correlation Coefficient	.476**	1.000	.233**
	Sig. (2-tailed)	0.000		0.000
	N	4338	4338	4332
Likelihood of Enrollment in Q1 or Q2	Correlation Coefficient	.368**	.233**	1.000
	Sig. (2-tailed)	0.000	0.000	
	N	4332	4332	4332
**. Correlation is significant at the 0.01 level (2-tailed).				

Table A7: Likelihood of Enrollment and Economic Interest Correlations

Correlation of Likelihood to enroll and Economic Interest			
Spearman's rho		Likelihood of Enrollment in Q1 or Q2	Student Economic Interest
Likelihood of Enrollment in Q1 or Q2	Correlation Coefficient	1.000	.397**
	Sig. (2-tailed)		0.000
	N	4332	4296
Student Economic Interest	Correlation Coefficient	.397**	1.000
	Sig. (2-tailed)	0.000	
	N	4296	4302
**. Correlation is significant at the 0.01 level (2-tailed).			

Appendix G: The Bootstrap Likelihood Ratio Test (BLRT)

The likelihood ratio test (LRT) is used to compare the fit of two models to determine if p-values indicate it is significant to increase c , the number of classes. However, this test allows for parameters to lie at 0 or 1, the boundaries of parameter space. The issue arises because this violates a regularity condition of defining parameters in terms of the subsuming model in LCA (Steiger et al., 1985; Holt and Macready, 1989; McLachlan and Peel, 2000; Nylund et al., 2007). Thus, we turn to the BLRT as a measure of model fit.

The BLRTs are used to measure fit and significance of additional classes in the model. It empirically estimates the log-likelihood difference using $-2LL$ to generate a p-value for comparison the distributions of the BLRT model and the true model.

Presently there has been an increase in the use of the bootstrap application due to user friendly software and increased efficiency. The constant innovation of technology and software (e.g. Latent Gold Choice, Mplus, among others) has led to decreased complexity and increased availability of more advance techniques for LCMs, such as the L^2 and BLRT methods.¹⁶ These techniques, mainly the BLRT, are rich in complex mathematics and have become more easily implementable in LCA due to software advances. In search for the correct number of classes or clusters in LCMs, the BLRT has shown success in consistently estimating the correct number C classes (see Nylund et al., 2007; Jung and Wickrama, 2008).

Dziak et al. (2014) explore this process extensively, concluding that the BLRT is the best LRT for LCA. They also provide technical formulas for deploying the BLRT method to determine class numbers (see Dziak et al., 2014). Another recent study uses the BLRT to illustrate and correct the effects of asymptotic distribution when estimating nuisance parameters (e.g. variance; σ^2) (see Jonker and Van der Vaart, 2014). Chan et al.

¹⁶ The L^2 test relaxes the assumption of chi-squared difference statistic used to estimate model fit to data. The L^2 LRT can be considered for model fit by estimating the p-values of the $c - 1$ and c class models to assess whether there is a statistically significant improvement by adding one additional class (see Nylund et al., 2007; Latent Gold Choice Technical Guide, 2005).

(2012) use the BLRT to measure their information criterion efficiency in a stochastic dominance approach to determine UK market efficiency. These studies exemplify the broadening popularity and the multi-functionality of using the BLRT.

The BLRT is used as follows. Consider this study where $C - 1$ is the null hypothesis and the 3-class model, the C is the 4-class model, and the alternative hypothesis. A rejection of the null hypothesis ($p < 0.05$) occurred in this case and indicates that the 3-class model is rejected for the 4-class model. The BLRT results were consistent with the L^2 results, suggesting a 4-class model will fit the data the best. The SABIC also indicated that the 4-class model was a better fit for the data than all the other model (see Section 4.3). See Holt and Macready (1989) for how p-values are computationally derived under this fit criterion. All bootstrap testing in this study was done with LGC.

Appendix H: Parameter Estimates

Tables A8 and A9 are parameter estimates used in the final analysis. Table A7 consists of attribute parameters and Table A8 consists of covariate estimates. Italicized and lower case attributes (e.g., *studyhr*) indicate the respective attribute and all caps (e.g., HIGH) indicates the various attribute levels.

Table A8: Final Model Parameter Estimates

4-Class Model for Choices								
	Class1	Class2	Class3	Class4	Overall			
R²	0.3995	0.1419	0.4528	0.4774	0.4248			
R²(0)	0.4082	0.1456	0.4788	0.4953	0.4249			
Attributes	Class1	Class2	Class3	Class4	Wald	p-value	Mean	Std.Dev.
<i>focus</i>								
RESOURCE USE/ ENERGY MARKETS	-2.0571	1.2022	-3.1962	7.809	27.1078	0.0017	-0.0927	3.4012
CONSERVATION/ SUSTAINABILITY	-2.3994	0.5638	-1.891	12.8468			0.4522	4.8319
BALANCE USE AND CONSERVATION	0	0	0	0			0	0
<i>major</i>								
GRADUATION ONLY	-1.2961	-0.3836	-0.9671	2.9897	43.5616	0.026	-0.4369	1.3439
MAJOR REQUIREMENT FULFILLED	0	0	0	0			0	0
<i>studyhr</i>								
HIGH (7HRS)	-1.6259	-0.1812	0.2068	3.9139	40.9667	7.40E-07	-0.1772	1.7157
MED (5HRS)	1.9956	-0.0127	-0.0849	5.8404			1.5023	1.8895
LOW (3HRS)	0	0	0	0			0	0
<i>profrank</i>								
INSRUCTOR/ LECTURER	-0.3689	-0.2581	3.1144	-6.4066	33.9131	0.00014	-0.4411	2.6028
JUNIOR PROFESSOR	-0.9871	-0.0879	2.4118	-9.681			-1.1751	3.4248
TENURED PROFESSOR	0	0	0	0			0	0
<i>smpaper</i>								
LARGE % OF GRADE	-2.326	0.0509	-3.3061	-0.3661	39.8083	2.80E-08	-1.5714	1.3162
SMALL % OF GRADE	0	0	0	0			0	0
<i>time</i>								
MORNING	0.0566	-1.2057	3.7481	-2.2075	53.0551	5.90E-05	0.0914	1.9039
MIDDAY	0.0538	-0.0976	2.3076	2.4941			0.7304	1.109
LATE AFTERNOON	0	0	0	0			0	0
<i>MWF</i>								
T/TH	-1.607	0.2912	0.9151	4.5988	20.2722	0.00024	0.1844	1.9502
MWF	0	0	0	0			0	0
<i>fall</i>								
SPRING	0.6093	0.2862	-3.1399	-3.6401	25.8395	0.04	-0.7074	1.7732
FALL	0	0	0	0			0	0

Table A9: Covariate Estimates

Intercept	Class1	Class2	Class3	Class4	Wald	p-value
	0.2185	0.5872	-0.3637	0	4.2082	0.24
Covariates	Class1	Class2	Class3	Class4	Wald	p-value
<i>INTEREST</i>						
LOW	-0.2317	-1.541	-0.6264	0	12.4682	0.052
MED	0.1717	-0.4712	0.3129	0		
HIGH	0	0	0	0		
<i>POLPREF</i>						
CON	2.0541	1.8467	1.783	0	15.9321	0.0012
LIB	0	0	0	0		
<i>USA</i>						
FOREIGN	1.5414	3.1892	-1.1637	0	9.6906	0.021
DOMESTIC	0	0	0	0		

Appendix I: Descriptive Statistics

Table A10 shows the descriptive statistics for the choice data. This was generated by SPSS. One thing to note in the last column is the “Valid N” which shows 90% of N had nothing missing (see Section 4.3). Table A11 shows LGC descriptive statistics of the final model.

Table A10: Descriptive Statistics

Descriptive Statistics					
	N	Minimum	Maximum	Mean	Std. Dev
Survey Version	4338	1	3	1.99	0.817
SET	4338	1	18	9.46	5.193
Awareness of NR Economics Courses	4338	1	5	2.08	1.185
Awareness of BS NRE Master’s Program	4338	1	5	1.6	0.981
Likelihood of Enrollment in Q1 or Q2	4332	1	5	2.09	1.246
Choice Set	4324	1	2	1.49	0.5
Student Rank	4296	0	2	0.82	0.426
Student Sex	4302	0	1	0.77	0.419
Students Citizenship Status	4266	0	1	0.96	0.198
Student Major Focus	4308	0	2	1.69	0.682
Student Economic Interest	4302	0	2	0.96	0.613
Students Post Graduation Plans	4248	0	2	1.73	0.524
Students Job Preference	4254	0	2	0.77	0.593
Students Political Preference	4098	0	1	0.4	0.491
Valid N (listwise)	3946				

Table A11: LGC Model Description Statistics

4-Class Choice Model		
Number of cases	675	
Number of replications	4042	
Number of parameters (Npar)	63	
Random Seed	549699	
Best Start Seed	549699	
Chi-squared Statistics		
Degrees of freedom (df)	612	p-value
L-squared (L^2)	1006.925	9.10E-22
X-squared	1900.964	5.70E-132
Log-likelihood Statistics		
Log-likelihood (LL)	-2405.31	
SABIC (based on LL)	5021.011	
Classification Statistics	Classes	
Classification errors	0.1959	
Reduction of errors (Lambda)	0.6741	
Entropy R-squared	0.6371	
Standard R-squared	0.6081	

Appendix J: Response Frequencies

Summary of Survey Frequencies

The table below summarizes the response frequencies of the sample.

Table A12: Frequency Statistics

Frequency Statistics								
		Survey Version	SET	Awareness of NR Economics Courses	Awareness of BS NRE Master's Program	Likelihood of Enrollment in Q1 or Q2	Choice Set	Student Rank
N	Valid	4338	4338	4338	4338	4332	4324	4296
	Missing	0	0	0	0	6	14	42
		Student Sex	Students Citizenship Status	Student Major Focus	Student Economic Interest	Students Post Graduation Plans	Students Job Preference	Students Political Preference
N	Valid	4302	4266	4308	4302	4248	4254	4098
	Missing	36	72	30	36	90	84	240

NRE Awareness Questions

The following three tables display the response frequencies of the first 3 Likert-scale awareness questions.

Table A13: Awareness Question 1 Frequencies

Awareness of NR Economics Courses				
Awareness Level	Frequency	Percent	Valid Percent	Cumulative Percent
NOT AT ALL AWARE	1830	42.2	42.2	42.2
A LITTLE AWARE	1134	26.1	26.1	68.3
SOMEWHAT AWARE	792	18.3	18.3	86.6
VERY AWARE	348	8	8	94.6
EXTREMELY AWARE	234	5.4	5.4	100
Total	4338	100	100	

Table A14: Awareness Question 2 Frequencies

Awareness of BS NRE Master's Program				
Awareness Level	Frequency	Percent	Valid Percent	Cumulative Percent
NOT AT ALL AWARE	2796	64.5	64.5	64.5
A LITTLE AWARE	876	20.2	20.2	84.6
SOMEWHAT AWARE	396	9.1	9.1	93.8
VERY AWARE	150	3.5	3.5	97.2
EXTREMELY AWARE	120	2.8	2.8	100
Total	4338	100	100	

Table A15: Likelihood of Enrollment Frequencies

Likelihood of Enrollment in Q1 or Q2				
Likelihood	Frequency	Percent	Valid Percent	Cumulative Percent
NOT AT ALL LIKELY	1819	41.9	42	42
A LITTLE LIKELY	1266	29.2	29.2	71.2
SOMEWHAT LIKELY	666	15.4	15.4	86.6
VERY LIKELY	191	4.4	4.4	91
EXTREMELY LIKELY	390	9	9	100
Total	4332	99.9	100	
Missing	6	0.1		
Total	4338	100		

Attribute Response Frequencies

This table displays all the response frequencies for all of the sociodemographic questions asked.

Table A16: Attribute Response Frequencies

Student Rank					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	LOWER	834	19.2	19.4	19.4
	UPPER	3384	78	78.8	98.2
	GRAD	78	1.8	1.8	100
	Total	4296	99	100	
Missing	-88	42	1		
Total		4338	100		
Student Sex					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	FEMALE	978	22.5	22.7	22.7
	MALE	3324	76.6	77.3	100
	Total	4302	99.2	100	
Missing	-88	36	0.8		
Total		4338	100		
Students Citizenship Status					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	FOREIGN	174	4	4.1	4.1
	DOMESTIC	4092	94.3	95.9	100
	Total	4266	98.3	100	
Missing	-88	72	1.7		
Total		4338	100		
Student Major Focus					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	OTHER	540	12.4	12.5	12.5
	HASS	264	6.1	6.1	18.7
	STEM	3504	80.8	81.3	100
	Total	4308	99.3	100	
Missing	-88	30	0.7		
Total		4338	100		

Table A16 continued: Attribute Response Frequencies

Student Economic Interest					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	LOW	888	20.5	20.6	20.6
	MED	2682	61.8	62.3	83
	HIGH	732	16.9	17	100
	Total	4302	99.2	100	
Missing	-88	36	0.8		
Total		4338	100		
Students Post Graduation Plans					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	OTHER	168	3.9	4	4
	GSC	798	18.4	18.8	22.7
	JOB	3282	75.7	77.3	100
	Total	4248	97.9	100	
Missing	-88	90	2.1		
Total		4338	100		
Students Job Preference					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	OTHER	1362	31.4	32	32
	BUS	2526	58.2	59.4	91.4
	GOV	366	8.4	8.6	100
	Total	4254	98.1	100	
Missing	-88	84	1.9		
Total		4338	100		
Students Political Preference					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	CON	2442	56.3	59.6	59.6
	LIB	1656	38.2	40.4	100
	Total	4098	94.5	100	
Missing	-88	240	5.5		
Total		4338	100		

Appendix K: Crosstabs Analysis

This sections consists of cross tabulation analysis of the survey awareness questions and a couple different significant sociodemographic characteristics (citizenship status, economic interest).

Table A17 supports the assertion that 0 foreign students stated they were “extremely aware” of the NRE master’s program. Evidence that the awareness of these particular students is low and there should be some effort to reach these students.

Table A18 shows the relationship between student’s likelihood of enrollment and their economic interest level.

Table A17: Awareness of NRE Master’s Program & Students Citizenship Status

		Citizenship Status		Total
		FOREIGN	DOMESTIC	
Awareness	NOT AT ALL AWARE	66	2700	2766
	A LITTLE AWARE	72	792	864
	SOMEWHAT AWARE	24	348	372
	VERY AWARE	12	132	144
	EXTREMELY AWARE	0	120	120
Total		174	4092	4266

Table A18: Likelihood of Enrollment in Q1 or Q2 & Student Economic Interest

		Student Economic Interest			Total
		LOW	MED	HIGH	
Likelihood of Enrollment in Q1 or Q2	NOT AT ALL LIKELY	588	1135	90	1813
	A LITTLE LIKELY	210	840	198	1248
	SOMEWHAT LIKELY	66	462	132	660
	VERY LIKELY	12	95	78	185
	EXTREMELY LIKELY	12	150	228	390
Total		888	2682	726	4296

Appendix L: Sampled Courses Summary

The table shown below indicates the courses surveyed, their professors, and their respective sample contribution.

Table A19: Sampled Courses Summary

Professor Name	Course Name	Course #	Frequency
Dr. William S. Breffle	Environmental Economics	EC 4650	11
Dr. William S. Breffle	Environmental Economics (Graduate Level)	EC 5650	3
Dr. Paul A. Nelson	Senior Seminar in Economics	EC 4000	4
Dr. William S. Breffle	Economic Decision Analysis	EC 3400	54
Dr. Gary A. Campbell	Mineral Industry Economics	EC 4630	17
Dr. Gary A. Campbell	Mineral Industry Economics (Graduate Level)	EC 5630	2
Dr. Gary A. Campbell	Principles of Economics	EC 2001 B	70
Dr. Paul A. Nelson	Industrial Organization	EC 3300	53
Bryan Lagalo	Economic Decision Analysis	EC 3400 (A)	131
Bryan Lagalo	Economic Decision Analysis	EC 3400 (B)	120
Dr. Emanuel Xavier-Oliveira	International Economics	EC 3100 (A)	20
Dr. Emanuel Xavier-Oliveira	International Economics	EC 3100 (B)	55
Dr. Latika Gupta	Principles of Economics	EC 2001 (A)	106
Bryan Lagalo	Microeconomic Theory	EC 3002	55
Dr. Latika Gupta	Energy Economics (Undergrad & Grad.)	EC 4620/5620	22
		Total	723
*The prefix EC simply denotes that it is an economics course.			