ABSTRACT

Title of Dissertation: **DYNAMIC STRATEGY CHOICE BEHAVIOR UNDER COMPETITIVE ENVIRONMENT: APPLICATION TO ELECTRONIC FREIGHT AUCTION MARKETPLACES**

Yeonjoo Min, Doctor of Philosophy, 2007

Dissertation directed by: **Professor Hani S. Mahmassani**
Department of Civil and Environmental Engineering

In the electronic market place, the auction is known to be economically efficient, allowing players to maximize their own benefits. Through this mechanism, new spot markets are created, which connect buyers and sellers. In this new spot markets, many problem contexts give rise to competitive decision situations in which players must make repeated decisions along with or in response to competitors' decisions. Auction-based electronic marketplaces for freight service procurement are an example of such environments, and provide the motivating application context for the models presented in this paper.

The specific focus is on the decisions of carriers, as bidders for the loads tendered by shippers in spot market situations. This paper is about the learning models used to describe a player's strategy choice behavior using experimental data and explains how that choice arises from the nature of multi-player interactions and their dynamics over multiple bids. Therefore, the principal focus of this paper is how to model a player's dynamic strategy choice behavior under the pressure of competition.

A dynamic strategy choice model structure for two type of cognitive learning process is formulated, with alternative specifications corresponding to different levels of cognition capacity. Furthermore, the dynamic strategy choice model structure for mixed learning is developed, which combines both elements of two different learning processes. The model is intended to describe how a player or agent in a non-cooperative game with no perfect information and bounded rationality chooses a bidding strategy. We propose a general dynamic strategy choice model framework using the dynamic mixed logit model structure and estimate the model parametrically, using two sets of experimental data. The paper also presents econometric issues that arise in estimating such models given a time series of auction bids and outcomes, and formulates error structures appropriate to the highly interactive dynamic nature of competitive auction-based marketplaces.

DYNAMIC STRATEGY CHOICE BEHAVIOR UNDER COMPETITIVE ENVIRONMENT: APPLICATION TO ELECTRONIC FREIGHT AUCTION MARKETPLACES

By

Yeonjoo Min

Dissertation submitted to the Faculty of the Graduate School of

University of Maryland, College Park in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

2007

Advisory Committee:

Professor Hani S. Mahmassani, Chair
Professor Ali Haghani
Professor Michael Ball
Professor Paul Schonfeld
Assistant Professor Cinzia Cirillo

Table of Contents

List of Tables

List of Figures

Chapter 1. Introduction

1.1. Motivation

The booming world of electronic commerce now provides bidders using electronic auction systems with virtual agents to do their bidding. In the electronic market place, the auction is known to be a well-established and powerful social information-processing mechanism. In addition, it is well known to be economically efficient, allowing players to maximize their own benefits. Through this mechanism, new spot markets are created, which connect buyers and sellers. Electronic market places offer benefits to both buyers and sellers by reducing transaction times, costs and effort. Buyers can search for and compare providers easily in a marketplace; while, for sellers, marketplaces provide access to broad customer bases. However, a complicated decision-making process has emerged in the electronic market place, wherein auction systems now generate many benefits over traditional electronic market places.

In particular, many problem contexts give rise to competitive decision situations in which players must make repeated decisions along with or in response to competitors' decisions. The motivating problem context for this dissertation consists of auction-based electronic marketplaces for freight service procurement. These virtual environments give rise to new classes of decision situations for carriers and shippers. The specific example that provides the motivation and the focus for the present study are the decisions of carriers, as bidders for the loads tendered by shippers in spot market situations.

Game theory has been the primary theoretical approach for those seeking to understand how markets might evolve under different assumptions on player (carriers in this case) behavior. It has placed considerable emphasis on equilibrium, and understanding the extent to which equilibria may be reached in markets under different information situations and payoff structures. However, most game theoretic constructs assume that the players (agents) have complete[1] or perfect[2] information, common knowledge[3] and perfect rationality[4]. In practice, these assumptions are too restrictive for most real-world situations, especially in fast-moving dynamic environments with repeated auction plays. Relaxation of these assumptions places greater emphasis on the cognitive processes and individual characteristics of players in analyzing the associated decision-making processes.

Choice behavior in non-cooperative auction games, under competitive environments, differs in several respects from the general travel choice behavior

---

[1] Each player is aware of all other players, the timing of the game, and the set of strategies and payoffs for each player.
(source: http://www.gametheory.net/Dictionary/)

[2] A sequential game is one of perfect information if only one player moves at a time and if each player knows every action of the players that moved before him at every point. (source: http://www.gametheory.net/Dictionary/)

[3] An item of information in a game is common knowledge if all of the players know it and all of the players know that all other players know it, and so on. (source: http://www.gametheory.net/Dictionary/)

[4] Players always act in a rational way, and are capable of arbitrarily complex deductions towards that end.
(source: http://en.wikipedia.org/ wiki/)

typically considered in transportation demand studies, e.g. to represent the traveler's choice of mode, departure time, residential or other location, etc. In this case, the other players', as well the player's own, previous decisions and payoffs affect current choice and payoff. The high degree of interaction among competitors is a critical phenomenon in dynamic game situations. Furthermore, players in a game can communicate with each other by exchanging information about strategies and payoffs, either explicitly or implicitly, through imitation and adaptation of successful strategies. However, typical discrete choice model formulations for everyday transportation demand situations do not include the impact of such interaction among decision makers in their respective choice behavior, because competition is not a consideration in these decision situations. In dynamic auction games, one's strategy depends on whether other players are bidding their value or are shading their bids. Players can learn how to play those games over time, and can update their belief by learning how to play based on their past experience and acquired knowledge of the opponents' actions and their payoffs. However, that type of information is limited to the type of auction and public information disclosed by auctioneers. This dissertation uses concepts from both discrete choice models and game theory to develop descriptive dynamic strategy choice models in competitive environments.

1.2. Research Objectives

The main objective of this research has been to develop a theoretical framework and methodology to model a player's dynamic strategy choice behavior within a competitive environment, by extending and adapting the existing model framework. This dissertation explores how bidders compete in auction-based electronic marketplaces for freight-service procurement. The study considers the carrier's choice of strategy of how much to bid to acquire a tendered load in an auction-based marketplace with repeated auctions involving the same set of players. A dynamic strategy choice model structure for two type of cognitive learning process is formulated, with alternative specifications corresponding to different levels of cognition capacity. Furthermore, the dynamic strategy choice model structure for mixed learning is developed, which combines both elements of two different learning processes. The model is intended to describe how a player or agent in a non-cooperative game with no perfect information and bounded rationality chooses a bidding strategy. The study also addresses econometric issues that arise in estimating such models given a time series of auction bids and outcomes, and formulates error structures appropriate to the highly interactive dynamic nature of competitive auction-based marketplaces. The model structure and estimation process are illustrated using a data set obtained from experiments conducted with players in hypothetical bidding situations.

The specific goals of this research are as follows:

1. To state and formulate a player's cognitive-process behavior and his or her dynamics in bidding behavior under two different assumptions: i) the availability of feedback information; and ii) the level of cognitive and instrumental rationality.

2. To develop mathematical models with which to capture a player's dynamic competitive behavior relating to bidding price choice under different assumptions of that player's level of cognition capability. More specifically,

    a. To provide the mathematical framework for the dynamic multinomial probit model (DMNP)

    b. To provide the mathematical framework for the dynamic mixed (or kernel) logit model (DML)

The models described in this dissertation extend previous work (Lam 1991; Liu 1997; Srinivasan 2000), by incorporating a model framework for a player's competitive bid-price choice behavior. A new error structure is explored to capture players' interactions during their respective choice behaviors, which is different from the error structure of the existing model framework. Furthermore, the joint choice probability function among players is applied to estimate parameters in learning models, and it is formulated differently for both the dynamic multinomial probit and the dynamic mixed logit models.

3. To present the different model specifications of the epistemic and behavioral reinforcement learning processes.

Each cognitive learning model specification is different, depending upon the different levels of player cognition capacity and the limit of feedback information.

4. To present the model specifications for a mixed-learning process, so as to provide a general form of model specification that incorporates both the epistemic and behavioral reinforcement learning models.

The processes can be mixed, if associated with different periods, players or mechanisms, and deepened by incorporating the reasoning principles.

5. To estimate parameters for a dynamic player's bid-price choice behavior model, using experiment panel data in hypothetical bidding situations

Even though the data are obtained from experiments in which decision-makers are in hypothetical bidding situations, the data are useful for developing insights into the underlying players' strategy choice behavioral processes, and into their decision-making interactions during choice behaviors.

6. To comparatively analyze and interpret a player's strategy choice behavior between different cognitive learning processes.

The explanatory analysis results generated for two cognitive learning processes and using two types of experimental data are compared, so as to provide insights into each player's underlying behaviors. The estimation results are interpreted and compared to explain players' strategy choice behavior within the context of each learning model. Behavioral patterns generated by the two cognitive learning processes also are compared. In addition, the possible transferability of behavioral insights and models across learning processes are examined.

1.3. Research Approach

This research starts by studying the cognitive dynamics of four learning processes, ordered in terms of a player's decreasing cognitive capacities. Two intermediate learning processes, epistemic and behavioral reinforcement, are explored, while two extreme processes, eductive and evolutionary, are excluded in this dissertation.

Based upon these cognitive learning concepts, the mathematical frameworks are formulated to incorporate a player's competitive behavior into a strategic choice decision-making process. Also, two experiments including two types of game are conducted to collect datasets, including observations of player's bid price decisions in repeated auction games. The dynamic strategy choice models are specified and estimated using experimental data and model frameworks associated with the different types of cognitive learning behavior. More specifically, the major tasks of this dissertation are:

1. To develop a framework to incorporate a player's interaction effect into the decision-making process by which a player changes his or her strategy for winning the next game and maximizing personal payoffs.

The model framework assumes that players change their strategy based upon their past experiences or the beliefs about what other players will do.

2. To conduct an experimental survey and collect data to observe players' choice-decision behaviors under conditions in which different levels of feedback information are available to them.

Those experiments include a player's bidding history in repeated games, including past bidding price decisions.

3. To specify the dynamic strategy choice models for epistemic and behavioral reinforcement learning processes.

The components of these models are derived from the first task by developing a framework for the cognitive learning process. The model specifications are different, depending upon the assumptions behind each cognitive learning process.

4. To develop the mixed type of dynamic strategy choice model for use formulating the general model specifications that both epistemic and behavioral reinforcement learning processes might explain.

This mixed learning model can combine the elements of two seemingly-different approaches and include them as a special case.

5. To formulate error structures appropriate for the highly-interactive dynamic nature of competitive auction-based marketplaces, and to specify the joint probability function over players.

6. To estimate the parameters in the dynamic strategy choice model specification system used for epistemic and behavioral reinforcement learning processes, using experimental data in hypothetical bidding situations.

Estimated parameters capture the effect of a player's past experiences and his or her average payoff in a behavioral reinforcement learning model. They also capture the effect of a player's possible payoff based upon beliefs regarding the type of opponent they are facing in epistemic learning model. Moreover, the unobserved influences of a player's competitive interaction on the decision-making process are captured in error structures. Both models include habit persistence terms, which capture the effect of previous utility and unobserved serial correlations on the current player's strategy choice decision, over various time periods (or games).

7. To examine the transferability of a player's bidding behavior between epistemic and behavioral-reinforcement learning models, and to compare the substantive insights generated from each of these two 'pure' cognitive learning models to those generated using a mixed (hybrid) dynamic strategy choice model.

8. To estimate the parameters for the mixed learning dynamic strategy choice models using experimental data.

The estimation results indicate the effect of a player's past experiences and his or her average payoff in a behavioral reinforcement learning model and the effect of a player's possible payoff based upon beliefs regarding the type of opponent they are facing in epistemic learning model. In addition, estimated parameters show the player's propensity of each cognitive learning process and capture the effect of the interaction among players' cognitive learning rules.

9. To interpret a player's choice-decision behavior for each cognitive learning process. To compare the results from both cognitive learning processes and

explain how players behave differently in response to limited feedback information.

The present dissertation essentially considers how well simple learning models, motivated by the psychology of learning, can model who must learn about the game and who must learn about their opponent during the course of playing a game, over time. Our goal is to model observed behavior, starting with behavior that is observed in experimental settings. In conclusion, we also consider the implications of this approach for applied economics in naturally-occurring, non-experimental settings. We show that experimental data can be both well-described and robustly-predicted by relatively simple learning theories.

The following Figure 1.1 shows the research approach structure.

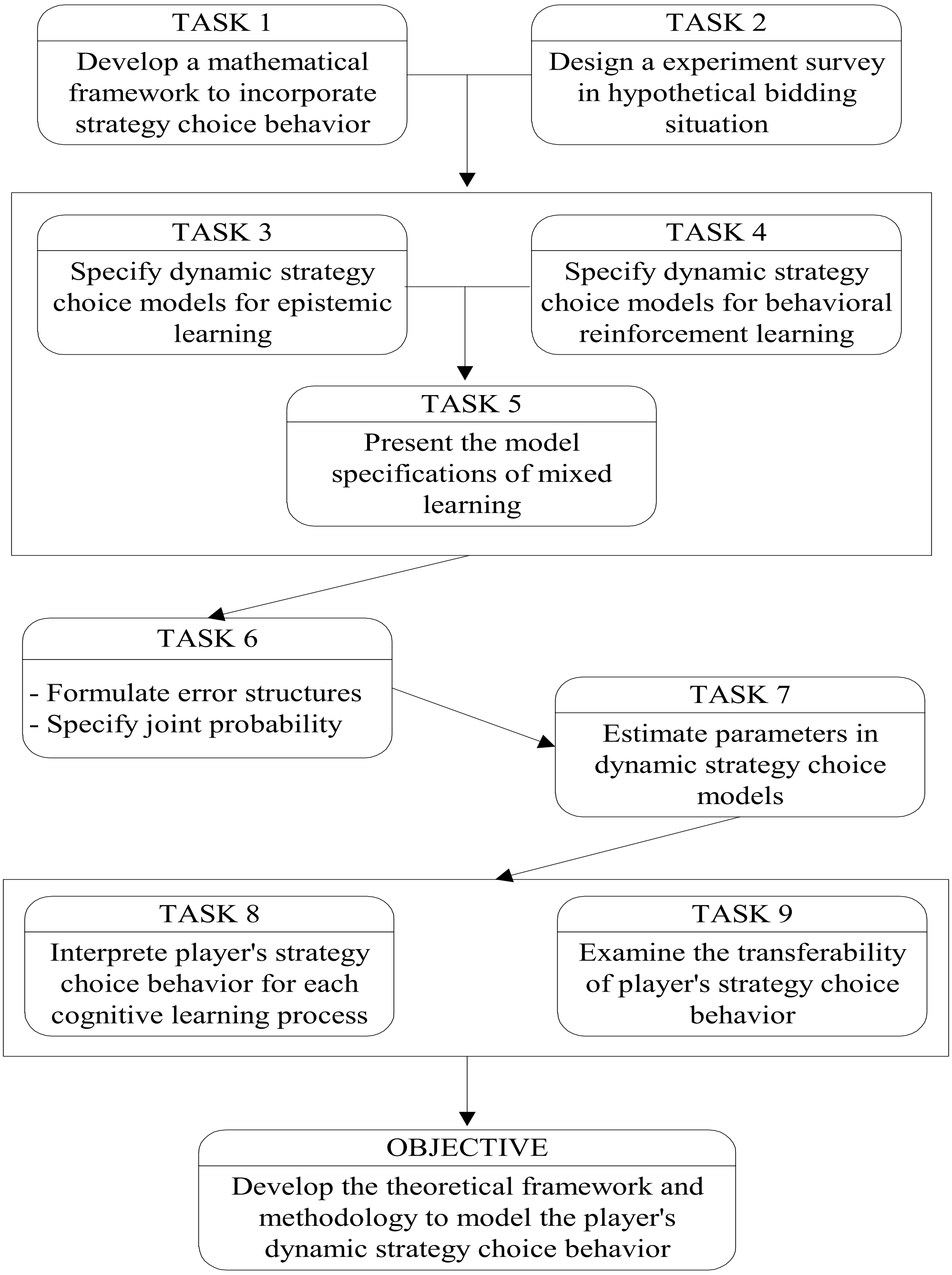


Figure 1.1 Study Approach

1.4. <u>Organization of the Dissertation</u>

The organization of this dissertation proposal is as follows. Already, the introduction chapter has provided a brief overview, including problem definitions, the motivation behind the project, the approach and the objectives of this dissertation. In Chapter 2, the four cognitive learning processes are reviewed. These processes are: 1) eductive learning; 2) epistemic learning; 3) behavioral reinforcement learning; and 4) evolutionary learning. Chapter 3 summarizes the literatures and, while doing so, discuss five related topics: 1) Competitive Environments in Auctions; 2) Nash Equilibrium vs. Probabilistic Equilibrium: Noisy Behavior in Auctions; 3) Game Theory vs. Discrete Choice Modeling; 4) Game Theoretic Learning Models; and 5) Choice Models.

Chapter 4 presents the methodology used for the current research. The research methodology includes both the dynamic multinomial probit and dynamic mixed logit model frameworks for both the epistemic and behavioral learning processes. Chapter 5 presents the results of the explanatory analysis generated using the above-mentioned methodology, using two experimental datasets. Chapter 6 presents estimation results for the dynamic strategy choice model, using a dynamic multinomial probit estimation program for all of the cognitive learning processes, and then discusses these results. In Chapter 7, the dynamic mixed logit model estimation results for the dynamic strategy choice behavior are presented for both experiments. The last chapter provides a summary of and conclusions derived from this dissertation.

Chapter 2. Cognitive Learning Processes

2.1. Background Review

Modeling a player's choice of bidding strategy in an auction-type electronic marketplace, given observations of the sequences of actions and respective payoffs for multiple players, entails explicit assumptions about 1) the availability of information to the players, and 2) the associated cognitive-learning or adaptation process taking place in this environment. The adaptive rule considers each player's cognitive capacity, and results in certain model structures. This section presents the four cognitive learning processes principally recognized in the literature. The methodology section subsequently introduces two model specifications associated with two of these four learning processes, and introduces the framework used to estimate the corresponding dynamic strategy choice behavior models.

Each process is distinguished by the player's cognitive ability, which is broken down into two steps: 1) his/her available information to his intended strategy; and 2) two types of rationality: *cognitive* rationality and *instrumental* rationality (Walliser, 1998). Cognitive rationality deals with consistency between available information and constructed beliefs. Instrumental rationality entails consistency between given opportunities and fixed preferences, to determine strategies from prior expectations (Walliser, 1998). Each cognitive learning process has a different level of cognitive and instrumental rationality.

The four types of cognitive processes introduced by Walliser (1998) consist of the following, listed in decreasing order of a player's cognitive capacity: (1) eductive learning; (2) epistemic learning; (3) behavioral reinforcement learning; and (4) evolutionary learning. Figure 2.1 shows the decreasing order of cognitive capacity for each learning process. As noted, each type of cognitive learning process is differentiated from the others by the respective levels of cognitive and instrumental rationality. These can be described as follows:

- Eductive process: "each player has enough information to perfectly simulate the others' behavior and gets immediately to the equilibrium" (Walliser, 1998).

- Epistemic learning: "each player updates his beliefs about others' future strategies, with regard to their sequentially observed actions." (Walliser, 1998). (e.g. Belief Based Model)

- Behavioral reinforcement learning: "each player modifies his own strategies according to the observed payoffs obtained from his past actions." (Walliser, 1998).

- Evolutionary learning: "each agent has a fixed strategy and reproduces in proportion to the utilities obtained through stochastic interactions." (Walliser, 1998).

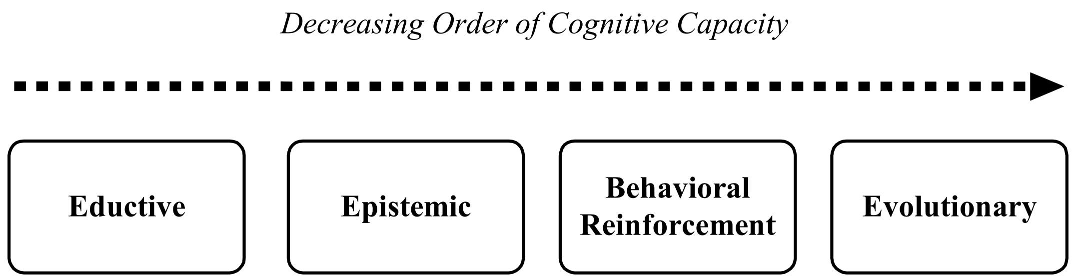


Figure 2.1 Order of cognitive rationality by learning process

This study excludes two extreme cases - eductive learning and evolutionary learning - because these two learning processes assume either perfect rationality or null rationality. In reality, a player has a limited ability to rationalize his or her decisions in an auction game. This study presents two dynamic strategy choice model structures, corresponding to the assumptions behind the following two types of cognitive learning process: epistemic learning and behavior-reinforcement learning.

2.2. <u>Eductive Learning Process</u>

The eductive learning process (Binmore, 1987), is generally used for a one-shot game; it assumes that a player has complete prior information about the other players' rationality and about their characteristics (opportunities, prior beliefs, and preferences) (Walliser, 1998). Game structure is assumed to be common knowledge, which means that players know others' knowledge about game. Therefore, a player with perfect cognitive rationality can simulate other players' behavior.

With eductive learning, each player enjoys perfect instrumental rationality and chooses a strategy by optimizing the utility function on his or her strategy set, corresponding to the opponents' expected strategies. Players can simulate the opponent's behavior, as well as their own behavior in a completely strategic situation. This strong assumption leads to pure Nash equilibrium. In the eductive process, a player's behavior is perfectly described and there are no variants, since the assumption is that the player's level of rationality is extremely high.

2.3. Epistemic Learning Process

Epistemic learning assumes that a repeated game is played over a sequence of periods. Players always know their own characteristics, but they have only incomplete information regarding their opponents' behavior or rationality, as they can observe the whole sequence of the others' decisions and draw the distribution of opponent's actions. A player cannot uncover the opponents' cognitive types from their choice decisions, but may simply assume their opponents' instrumental rationality, such as that reflected by some form of stationarity of opponents' behavior (Walliser, 1998). Epistemic learning behavior is observed easily in reality. With epistemic learning, a player's behavior must be different from that observed with eductive learning, because beliefs are updated with respect to time and chosen actions. However, the observation and computational costs are higher for epistemic versus eductive learning.

Limited cognitive rationality of players is assumed, since they determine their strategy through conjecture about their opponent's future strategy, based on an observation of his/her past actions. Accordingly, players choose a best response to their present conjecture regarding others' strategies each period. A player predicts the other's future actions as some combination of past occurrences. The important feature of this epistemic learning model is specifying prior beliefs and computing an initial value of preference, based on that prior belief.

2.4. Behavioral Reinforcement Learning Process

The behavioral reinforcement learning process is also used for repeated games. A player always knows his/her own opportunities, but he/she is no longer aware of the type of opponent or his/her past actions and payoffs. A player having weak cognitive rationality only knows his/her past experiences, and revises his/her experience according to the utility obtained from each strategy in past actions. With behavioral reinforcement learning, each player has limited rationality, so that he or she imitates past actions, which are reinforced by strategies that have succeeded and inhibited by strategies that have failed in past games. Moreover, each player may imitate the behavior of other players who have succeeded in the past, if their strategies and results are readily observable. Hence, the utility of a player is updated by observing his or her sequence of actions and corresponding payoffs.

Behavioral reinforcement learning behavior again takes place in real time, and always is achievable, since behavioral reinforcement learning can be applied easily from period to period (Walliser, 1998). In addition, the information and computation costs are lower than for epistemic learning, since only the player's own results are checked and memorized, and expectations about other players' actions are not included in the computational process. Specifying the player's own preference and computing the initial value of that preference comprise the crucial features of the behavioral reinforcement learning model.

In epistemic learning, players try to find the maximum payoff, based on their opponents' choice probability, but in this case, even if a player has information about his/her opponent's choice, he/she just observes and imitates the best response of the competitors.

2.5. <u>Evolutionary Learning Process</u>

An evolutionary process also assumes a repeated game. A player is no longer aware of the game structure and only has information about his/her actions if the strategies depend on it. With evolutionary learning, each player has no cognitive rationality. A player behaves in a preprogrammed, time-invariant and unconscious way to choose a strategy, and his or her action automatically is decided based upon behavior type: 'if context C, then action A' (Walliser, 1998). A player does not make his/her own decision, since that player's behavior is not determined at an individual level any more but at a population or subpopulation level. Also, each player has null instrumental rationality, since a player is not aware of his/her utility; only the modeler has that information. This learning process is a highly-adaptive model, since new information is used at a subpopulation level to identify a new direction associated with better utilities (Walliser, 1998).

2.6. Comparison

2.6.1. Information

As stated earlier, the four cognitive learning processes, from eductive to evolutionary learning, can be ordered by decreasing cognitive capacity and decreasing information about one's strategic environment. In eductive learning, a player has complete knowledge about his opponent's behavioral rules and beliefs, and players can forecast their own future strategies in all circumstances. In epistemic learning, the player has limited information about others' past actions and payoffs, but players are no longer aware of the opponent's behavior principles. A player can formulate expectations regarding their opponent's strategies. In behavioral learning, players can access information about their actions and payoffs by observing their successes and failures in the past. With the evolutionary process, the player eventually acquires information about their actions from subpopulation decisions, and he or she can choose to imitate those subpopulation decisions.

In addition, progressing from eductive to evolutionary learning, the decision-maker holds increasingly limited information about his own type, and is endowed with decreasing instrumental rationality. With the eductive process, the player maximizes his expected utility under known constraints, including information that perfectly predicts the opponent's strategies. In epistemic learning, the player considers a utility function limited to short-term effects, since he or she observes other opponent's decisions in the past and then decides his or her action depending on that information. In behavioral learning, the player simply adapts his or her behaviors to past utility, in a probabilistic way. In the evolutionary process, the player no longer improves his utility, but adapts via a subpopulation utility-increasing mechanism (Walliser, 1998).

2.6.2. Interaction

A player's decision is affected by interactions between competitors, and this causes stochastic decision-making behaviors. The stochastic elements occur from disturbances related to imperfect observations and incomplete modeling. The eductive process reduces stochastic elements by providing the possibility of mixed strategies, since a player can predict other players' strategies. In epistemic learning, players' beliefs regarding others' behaviors are affected by stochastic uncertainty (Goyal-Janssen, 1995). Players' observations of other's results are affected by sample bias in behavioral learning, because there is no information about the opponent's type. Also, the player's own behavior becomes probabilistic. In the evolutionary process, player's meetings to make subgroups are stochastically driven (Walliser, 1998).

2.6.3. Dynamics

In general, from eductive to evolutionary learning, the learning processes are less and less affected by the virtual dynamics of game simulation in a player's mind, and more and more affected by the real dynamics of game progress, as a result of the player's moves (Walliser, 1998). With eductive learning, there is no dynamic in outside reality, since the player completely simulates the player's behavior, with common knowledge in his mind, before making a decision in all games. In epistemic learning, the objective dynamic effects are more involved with the opponent's past actions, which are incorporated in the players' beliefs and expectations. Behavioral reinforcement learning shows that subjective dynamics affect a player's decision-making behavior and are included in the prediction of his or her own actions in the future, resulting from observing the player's own past actions. In evolutionary learning, the reproduction of the next population's distribution involves stochastic matches (as a subgroup) of players, which demonstrate the dynamics (Walliser, 1998). Therefore, evolutionary learning assumes that there is pure temporality without reasoning.

Chapter 3. Literature Review: Theoretical Principles

This chapter presents the characteristics of auction games and reviews the relevant literatures on game and choice theories. The dissertation deals with a sequential-decision type of game and related cognitive learning models for strategic and bounded rational players. We focus on the literatures regarding dynamic learning and decision choice models that have been studied fundamentally in game and choice theories. In particular, the game theory literature related to decision behaviors in sequentially-repeated games is reviewed. This chapter includes an overview of existing cognitive learning and dynamic choice model frameworks.

Section 1 describes the definition and concepts of auctions as pricing mechanisms and the competitive circumstances that exist in auction games. In Section 2, Nash equilibrium and probabilistic equilibrium are compared. In Section 3, game theory and choice theory are compared. Section 4 discusses the assumptions of cognitive-learning models, based on the game theory in experimental economics, and presents the existing model frameworks for cognitive learning processes. Section 5 presents two types of dynamic choice model, and discusses the assumptions and estimation procedures that exist in choice models

3.1. The Competitive Environment of Auctions

3.1.1. Auctions as a Price Decision Mechanism

Fixed-pricing or static time-differential pricing mechanisms are widely used in many existing electronic markets, because of their simplicity. However, in reality, there is variation in customers' demands over time. For this reason, those fixed pricing mechanisms are insufficient. Therefore, the importance issues of e-services and efficient pricing have emerged in existing electronic markets. Price is an important signal for controlling fair allocation of resources (Lee and Szymanski, 2005). However, people have difficulty making price decisions, because of the variant dynamics of pricing. An auction can be a solution. An auction is a continuously-adjustable dynamic pricing mechanism that adapts efficiently to changing market conditions.

Auctions are market institutions with an explicit set of rules determining resource allocation and prices, based on bids from market participants (McAffee and McMillan, 1987). The design of an auction and a specified set of rules determine the type of auction model, the outcomes of the auction, and the system by which bidding is conducted, how information is revealed, how communication is structured between buyers and sellers, and how allocations and payments are settled (Figliozzi, 2004).

This dissertation does not deal with the design of an auction; rather, it analyzes how the dynamic learning process is conducted, or how information is processed in an auction game. The study contained herein adopts existing standard auction mechanisms and analyzes the effect of the competitive environment of the auction on players' bidding decisions. The decision problem is strategic in nature, due to the interdependence of competitors' bids, costs and profits (Figliozzi, 2004). Therefore, the competitive player's behavior is analyzed in the dynamic strategy choice model context.

As mentioned before, auctions are mechanisms for determining the price of an object or a service in the presence of multiple bidders, and can be analyzed as a choice game. In electronic market environments, the use of an auction provides several benefits. First, auctions generally are easy to understand, and they are easy to access by both customers and service providers. Second, the rules and procedures of an auction system usually are easy to implement in automated electronic environments (Bichler, 2001). Third, an auction eliminates any need for defining complex dynamic pricing structures. Fourth, auctions support decentralized pricing and, therefore, avoid abusive market practices (McAfee and McMillan, 1997). Fifth, auction mechanisms are remarkable, since the auction form relies on modern social information processing technologies (Milgrom, 1998). Interactive network technology makes it easy for decision makers to access the auction mechanisms.

Furthermore, auction-based electronic marketplaces for the procurement of transportation services are believed to provide a high level of service for shippers, while controlling and reducing their operational costs. Carriers also benefit through easier access to the market, and via more efficient operations associated with fewer empty movements.

3.1.2. Bounded Rationality under Price Competition in a Market

In the classic model of price competition, named after Bertrand (1883), equilibrium, when at least two firms are in the market, exists when price is equal to marginal cost. In effect, each firm makes zero profits, even in a duopoly. Since observations from real markets are not matched with this result, this phenomenon is called the 'Bertrand Paradox'. It reports the dynamic results of markets in which participants compete for prices: the effect of changing the number of competitors on outcomes of the market. The next situation explains the Bertrand model, and demonstrates the unique circumstance of Nash equilibrium, in which both firms choose zero value.

It is assumed that, if more than two firms intend to compete in the market, at least two of them will choose zero in any equilibrium. This game can generate examples of common critiques of the Bertrand model and Nash equilibrium. In particular, it has been argued, among economists, that certain assumptions of the Bertrand model and Nash equilibrium are not realistic. They point out that the Bertrand paradox goes away if the assumption is relaxed, if goods are not assumed to be homogeneous, if capacity constraints are introduced, or if firms are allowed to compete repeatedly (Dufwenberg and Gneezy, 1998). Furthermore, firms may have incomplete information about payoffs and cost functions or demands. Economists suggest an explanation that relies on bounded rationality. This illustrates the effect of 'noise' on the variability of decision outcomes, when there is competition among players.

Dufwenberg and Gneezy (1998) proved that, if any firm in the market may bid differently from Nash equilibrium as the outcome of a Bertrand model, this explains why deviations from the Bertrand outcome depend upon competition between a number of agents. Based upon this theoretical prediction, all firms are supposed to submit the lowest possible bid, irrespective of how many agents are in attendance. However, when investigators have tested this model experimentally, they have discovered that, during the initial stage (the first game), competitors set prices higher than they would with Nash equilibrium. In subsequent rounds, if more than two competitors attend the game, the winning bids typically converge rather rapidly towards the theoretical prediction, in two out of three treatments. These experimental findings suggest that learning plays a role, since behavior tends not to be consistent across time in all games.

In reality, it seems unlikely that each agent is fully convinced that every other agent will behave in accordance with equilibrium, since they assume that competitors are fully rational and have perfect information. A little bit of irrationality can occur in the variability of decision behavior in a game, and this differs from the theoretical prediction generated by Nash equilibrium, even if a large enough number of competitors interact. Accordingly, the effect of competition among competitors is a critical factor to explain a player's bidding behaviors while decision making.

3.1.3. Effect of Multiple bidders on Bid Decisions in Online Auctions

An analysis of sequential auctions can be applied for the same good, ordered in time. Online sequential auctions are different from non-internet sequential auctions, in that the number of auctions is not fixed. Bidders enter those auctions at different times, and may have participated in different numbers of auctions. This uncertainty in bidding behavior is an interesting issue. Few studies have studied the theory of optimal bidding strategy in sequential auctions. It is very good model by which to show how uncertainty may affect bidders' bidding strategy.

Arora *et al* (2003) assume that uncertainty occurs because of a lack of information about the number of bidders. The bidder is assumed to know the distribution of the reservation values of the other bidders, but the actual values are private to each bidder. Under the certainty assumption, the optimal bidding price is independent of the strategies of other bidders during the first auction. These investigators proved that bidders would like to bid less than their reservation values in a first auction (a player submits the highest bid price win the game). If they have greater payoffs in the second auction, a lower bidding price is provided. When the number of bidders increases, each bidder faces more competition in the second auction, which makes the second auction less valuable to him or her. Hence, bidders increase their bids during a first auction (Arora *et al*, 2003).

If the number of bidders is unknown, the optimal bidding price is independent of the number of bidders participating in the first auction, and it is independent of the first auction bidding function. This implies that rational bidders will bid higher if they expect a larger number of bidders in a second auction (Arora *et. al*, 2003). Under risky or uncertain circumstances, bidders will bid less in the first auction, in order to win the game, if the second auction is anticipated to be more risky. Under conditions of bidding uncertainty, these investigators assumed that bidders only perceive the distribution functions of a number of bidders. This demonstrates the impact of the mean number of bidders, and its variability on a new bidder's strategy. Both the mean number of bidders and its variability in the subsequent second auction affect a new bidder's decision in the first auction. A larger mean number of bidders in the second auction means more competition, therefore less value to new bidders; hence, new bidders bid higher during the first auction. In on-line auctions, we are not able to observe the market participants' expected number of bidders, nor are we able to observe the perceived variance in the number of bidders.

It is critical, then, to distinguish what the difference is between online and offline sequential auctions. The number of bidders in online auctions is likely to vary stochastically. In on-line auctions, more uncertainty can exist than in non-online auctions, and this affects the competition between players. Those uncertainty factors should be modeled and included as components of the utility function, and as unobserved noise factors.

3.2. Nash Equilibrium vs. Probabilistic Equilibrium: Noisy Behavior in Auctions

This section introduces the probabilistic equilibrium models proposed by Andern, Goeree and Holt (1999), especially their *logit equilibrium model* of noisy behavior in auction-like games with a *'Traveler's Dilemma'* example. This illustrates why we need to study the dynamic learning process and why we need to develop a dynamic strategy choice model structure. Additionally, it suggests to which cases those suggested models are most appropriately applied.

Nash equilibrium (strategic equilibrium) is a list of strategies, one for each player, which has the property that no player unilaterally can change his strategy and acquire a better payoff (Osborne, 2004). Nash equilibrium in these types of games is insensitive to parameter changes. It cannot show the effect of randomness, called *'noise'*, that can occur, because of unobserved shocks in preference. More often, it appears when the observed payoffs become approximately equal. This randomness can be modeled using a probabilistic choice function, such as a logit or probit model. The probabilistic choice related to noisy behavior is introduced by being incorporated within these payoff asymmetry effects. A probabilistic choice function can be applied to the expected payoffs or to the utility, and this probability distribution satisfies a *'rational expectations'* consistency condition (Anderson, Goeree, and Holt, 1999).

A probabilistic choice rule is to specify the utility function with a stochastic component, reflecting unobserved factors. A player is assumed to choose decision j such as claim amounts (bid price) and certain kinds of strategy choice sets. If x_n^j denotes player n's decision j with expected payoff $\pi(x_n^j)$, the utility function is the following (Anderson, Goeree, and Holt, 1999):

$$U_n^j = \pi(x_n^j) + \mu\varepsilon_n^j \tag{3.1}$$

where

$\pi(x_n^j)$: player n's payoff by choosing decision j

μ: a positive *'error'* parameter

ε_n^j: the realization of a random variable

Here, if the coefficient is close to zero ($\mu = 0$), then the decision with higher values of expected payoff is selected. This result is matched with the expected result from Nash equilibrium.

In this case, this utility function includes only one parameter (payoff), but also an error term. In this dissertation, a sequential auction game is considered, and we develop different model specifications and error structures associated with the uncertainty of a player's behavior and the player's level of cognitive capacity, based upon that player's dynamic learning process.

A choice density is proportional to an exponential function of expected payoffs and it is expressed as (Anderson, Goeree, and Holt, 1999):

$$\Pr(D_i) = \frac{\exp(\pi_i^e / \mu)}{\sum_j \exp(\pi_j^e / \mu)} \quad \text{for } j = \text{the discrete choice sets} \tag{3.2}$$

If there is a continuum of alternatives, the form of choice density is following:

$$f(x) = \frac{\exp(\pi(x)/\mu)}{\int_{\underline{x}}^{\bar{x}} \exp(\pi(x)/\mu)dy} \tag{3.3}$$

$$\pi'(x)f(x) - \mu f'(x) = 0 \tag{3.4}$$

The above equation shows the differential equation in equilibrium choice density. This payoff function can vary with the type of auction. The choice probabilities are a smoothly-increasing function of expected payoffs, so that these probabilities are affected by the asymmetries in the cost of deviating from the payoff-maximizing decision (Anderson, Goeree, and Holt, 1999). The logit rule determines players' equilibrium distributions. This is known as a *logit equilibrium*. The logit equilibrium equation can be derived by understanding the learning and evolutionary dynamics that exist in games. In general, learning can be modeled in terms of beliefs about others' decisions in epistemic learning, and in terms of a player's own experienced payoff during behavioral reinforcement learning.

The traveler's dilemma is very good example of what the difference is between Nash equilibrium and logit equilibrium. It also explains why the logit equilibrium equation is more realistic than the Nash equilibrium concept. The following paragraph explains the situation that exists in traveler's dilemma, as described by Basu (1994):

"Two travelers returning home from a remote island discover that the identical antiques they bought have been lost in the airplane. The airline manager proposes the following scheme to bring out the value of the articles. The two travelers are instructed to independently submit compensation claims between $80 and $200. The airline will reimburse each traveler the minimum of the two claims. In addition, if the claims differ, a reward of $80 will be paid to the person making the smaller claim and a penalty of $80 deducted from the reimbursement for the larger claimant." (Basu, 1994)

In the traveler's dilemma game, the unique Nash equilibrium claims $80 for both players. However, it seems very unrealistic that any individual, no matter how rational, will submit an $80 claim.

Capra, Goeree, Gomez, and Holt (1997) collected laboratory datasets to test the above traveler's dilemma experiment. Their laboratory data revealed the frequency of claims and that the prediction of this model is sensitive to changes in penalty/reward. As the penalty increases, the probability increases that the claim is close to the unique Nash prediction. As the penalty decreases, the probability increases that the claim will be far from the unique Nash prediction ($80). For the low penalty case, Nash equilibrium cannot explain the player's strategy choice behavior; because there is more unobserved stochastic factors affecting the player's strategy choice decision. These examples demonstrate that the probabilistic equilibrium approach can be applied to a wide variety of interesting economic contexts. In this case, for both low and high-penalty cases, the probabilistic choice model can describe a players' behaviors and can statistically match the actual claims and frequency datasets.

Anderson, Goeree, and Holt (1999) proposed that logit equilibrium is a one-parameter generalization of Nash, by including the unobserved shock of preference as an error term. It can be evaluated using a maximum likelihood estimation of the laboratory data. Logit equilibrium can explain human decision behaviors for both one-shot and multiple-shot games. The Nash equilibrium value is probably close to the optimal value for the one-shot game case. In this experiment, the investigators believed that the error term explained the unobserved behavior. Moreover, the coefficient of the error parameter, μ, reflects human behavior, because no one can be perfectly rational and have full information. Higher values of the error parameter μ make choice probabilities less sensitive to expected payoffs. This means that we need to specify what other factors can affect the decision-making process in the game, and how differently we can model it. Therefore, we can add more specific public or private information parameters to the utility function, including the payoff function, in order to specify more unobserved factors.

Choice is stochastic, and the distribution of random variables determines the form of the choice probability. Therefore, if the random variable is normally distributed, the probit model can be applied; whereas the logit model is used for random variables with a Gummble distribution. In a simple case, Anderson *et al* (1999) applied the logit model with strong assumptions about the unobserved error term. The random components of the utilities of the different alternatives in the MNL model were assumed to be independent and identically distributed (IID), and to have a Gumbel distribution. Also, the *independence of irrelevant alternatives* (IIA) property holds in the logit model (Ben-Akiva and Lerman, 1985). The above previous researches could not consider the interactions among players, because of this strong assumption in error terms. We have come to realize the limitations of the MNL model; hence, the new dynamic multinomial and mixed logit models are applied in this dissertation.

*3.3. **Game Theory and Discrete Choice Modeling***

Game theory has developed to provide insight into the outcome of game situations under different behavioral assumptions regarding the players' preferences and decision rules, and various assumptions about information availability and other game settings. Game theory has provided a natural first approach to analyze auction-based electronic marketplaces that have been popularized through the Internet for a variety of general-purpose and specialized applications. Such marketplaces for freight service procurement allow carriers to interact dynamically with shippers in the competitive acquisition of shipments (loads), and the assignment of loads to carrier fleets. Carriers compete by bidding for shipments tendered by the shippers. Recent work by Figliozzi, Mahmassani and Jaillet (2003a, 2003b, 2005) has focused on the effect of dynamics on bidding behavior, and developed a model framework for examining the performance of such marketplaces in terms of carrier profit and shipper service levels. That work has also highlighted the limitations of classical game theory in capturing dynamic interaction effects on bidders' behavior in repeated auction games (Figliozzi, 2004).

The perspective adopted in this study is that of an analyst or observer (which may also be a competitor) seeking to predict the outcome of the bidding process followed by a player in a repeated auction game. Because real situations may depart significantly from the ideal conditions assumed in classical game theory, especially with respect to the dynamics of information in repeated games, we present a dynamic strategy choice model framework that recognizes, in specification and parameterization, the nature of multi-player interaction and its dynamics over multiple bids (plays). Given actual observations of carrier bids under specific information availability scenarios, the model can be calibrated to reflect the particular interaction patterns present in that situation. The dynamic strategy choice model presented in this study uses both concepts from discrete choice theory and game theory to model bidder (carrier) behavior in repeated auction-type games. In order to understand the nature of the dynamic strategy choice model, this section describes the differences and similarities between game theory and discrete choice models.

Discrete choice models are commonly used in transportation demand studies, marketing and other disciplines to represent and predict decisions made by individuals facing discrete choice alternatives. The most commonly used model forms, such as generalized extreme value models, multinomial probit and mixed logit, are derived from random utility maximization (Ben-akiva and Lerman, 1987; Train, 2003). Decision-makers are assumed to base their choice on the relative utility they associate with each of the available alternatives. This utility is a latent variable, which depends in a systematic manner on various observable attributes of the alternative and the decision-maker, as well as on unobservable component which may only be known in distribution to an analyst interested in representing the decision-maker's preferences and choice process. The form of the distribution of unobservables and its properties determines the mathematical form of the choice probability function. In typical travel decision situations addressed through discrete choice models, the decision maker's evaluation of the alternatives and subsequent actions are not generally directly affected by other individuals' choices or preferences; exceptions include instances of household interactions and/or firm level decisions which take place in a cooperative setting. In static applications of discrete choice models, the individual's decision does not directly interact with another player's action or payoff. Even applications to dynamic decision settings have generally only considered the individual's own previous decisions and associated consequences (which may nonetheless in turn depend on the users' collective decisions). Under competitive circumstances, this is not a realistic mechanism to apply.

In game theory, players choose the strategy or action that maximizes their respective payoff. As is common in microeconomic theories, game theory assumes that players are perfectly rational, with common knowledge and unbounded computational capabilities. The well-known Nash equilibrium theory is also built on those strong assumptions. Under a Nash equilibrium (strategic equilibrium), no player can attain a better payoff by unilaterally changing his/her strategy (Osborne, 2004). A Nash equilibrium in these types of games is insensitive to parameter changes, both systematic and random, e.g. unobserved "shocks" in preferences (Anderson, Goeree, and Holt, 1999). Experimental evidence suggests that the final equilibrium predicted by this theory does not match observed equilibrium conditions too well. Since people learn what other people will do, and tend to adjust to it, experimental equilibrium values will often deviate from those expected under Nash equilibrium. Furthermore, people with different histories and characteristics do not typically attain the same common knowledge. Some of the unobserved variation across players could be modeled using a probabilistic choice function. However, to complete the representation in a repeated game setting, it is necessary to introduce a description of the dynamic (cognitive) learning process associated with human choice behavior.

Game theoretic formulations do not typically include latent variables in the model structure to describe individuals' (or agents') decision-making processes, under the assumption of perfect rationality. In actual games, a player's decisions will be affected by previous payoffs, as well as by current and past actions of competing players. The assumptions of perfect rationality and perfect knowledge or information under such conditions become difficult to sustain, and bounded rationality becomes a more plausible notion in describing the decision-making process and its dependence on individual history. If the player can predict his/her opponent's action in a game, he/she can choose the best action corresponding to the other player's action. Therefore, a player tries to (implicitly) model the opponent's behavior function based on historical data. The existing model framework from classical game theory is therefore not sufficient to describe and predict choice behavior under a competitive environment, such as repeated auction game situations encountered in electronic freight marketplaces. A dynamic strategy choice model is presented in this dissertation as a framework to describe players' choice behavior in such environment.

3.4. Game Theoretic Learning Models

3.4.1. Introduction

Since the 1950's, *game theory* has became a popular research field among economists. Game theory traditionally has been considered to be the theory of strategic interactions among players who are perfectly rational, and who exhibit equilibrium behavior (Erev, 1998). Game theory, as a part of economic theory, has exerted the greatest contribution to design auction mechanisms. Nash equilibrium has proven to be a powerful instrument to analyze a player's perfectly rational behavior. As mentioned before, in reality, people do not act in the way postulated by Nash equilibrium models. Since the late 1980's, *evolutionary game theory* has been studied as a means to explain a player's behaviors in an auction game. This section presents the existing mathematical frameworks in evolutionary game theory, and the learning assumptions that exist in experimental evolutionary learning models.

Several mechanisms have been proposed to explain player's choice behaviors in sequential repeated games. However, most of these perspectives have not been able to completely explain the actual time-scale of equilibration in complex games. In a non-cooperative game, people learn how to play through experience. Consequently, evolutionary learning theory is important to understanding equilibration theoretically, and to explaining the changes in strategic behavior observed in the lab and in the field.

In recent years, evolutionary learning theory has been proposed by game theorists and experimental economists. They have focused a great deal of their attention on the question: how do people learn in repeated games (Nyarko and Schotter, 2002)? The goal of this theory is to understand how equilibrium can arise in the long term for multiple players who need not to be rational or even conscious decision makers. The learning models consider the limitedly-rational player's adaptive behavior in a game. Experimental economists have modeled empirically-observed behavior, based upon the foundations of equilibrium theories.

Arthur (1991), Mookherjee and Sopher (1994), Roth and Erev (1995, 1998), and Borgers and Sarin (2000) proposed reinforcement learning models in which people learn by looking back at their experiences and seeing what has been successful for them in the past. Belief-based models in epistemic learning have been developed by several economists: Boylan and El-Gamal (1993), Mookherjee and Sopher (1994, 1997), Cheung and Friedman (1997), Rankin, Van Huyck, and Battalio (1997), Fudenberg and Levine (1998), and Nyarko and Schotter (2002). They focused on a player's epistemic learning behavior and assumed that beliefs are updated by their opponent's future actions. Furthermore, Camerer and Ho (1998, 1999, 2002) offered a mixed model based upon both reinforcement and belief-based models. In the next sections, those experimental learning models are reviewed.

3.4.2. Belief-based Models

Belief-based models are based on the concept of epistemic learning. Belief-based models assume that, while past actions and payoffs are observable, beliefs are unobservable and, therefore, must be represented by proxies and inferred (Nyarko and Schotter, 2002). In a γ-weighted belief model, the weighted average of past action is taken as a proxy of beliefs, and all weights are decreased by the ratio of γ (Rankin, Van Huyck, and Battalio, 1997).

Boylan and El-gamal (1993) proposed two belief-based models: Cournot[5] play and fictitious[6] play. They then compared the predictions generated by these two models. They identified which types of game were consistent with the Cournot and which with the fictitious play model. Cheung and Friedman (1997) also developed two types

---

[5] The Cournot adjustment model, first proposed by Augustin Cournot (1838) in the context of a duopoly, has players select strategies sequentially. In each period, an agent selects the action that is its best response to the action chosen by the competing agents in the previous period. Cournot learning can be the extreme form of fictitious play in which each firm assumes that its competitor is using the same strategy in every period which is equivalent to the one most recently used. (source: http://www.gametheory.net/Dictionary/)

[6] A process by which players assume that the strategies of their opponents are randomly chosen from some unknown stationary distribution. In each period, a player selects her best response to the historical frequency of actions of his/her opponents. (source: http://www.gametheory.net/Dictionary/)

of play, and compared their two different belief-based models. In addition, they developed a more general model, a form of hybrid model, which includes both types of play. The γ-weighted empirical belief in a hybrid model is defined by Cheung and Friedman (1997):

$$B_n^j(t) = \frac{I_n^t(a^j) + \sum_{u=1}^{t-1} \gamma_n^u I_n^{t-u}(a^j)}{1 + \sum_{u=1}^{t-1} \gamma_n^u} \qquad (3.5)$$

where

$B_n^j(t)$: player n 's belief about the likelihood that opponent choose action a^j in period $t+1$

$$I_n^t(a^j) = \begin{cases} 1 & \text{if action } a^j \text{ was chosen in period } t \\ 0 & \textit{otherwise} \end{cases}$$

γ_n^u : the weight given to the observation of action a^j in period u

Fictitious play is the special case, in which γ=1; and Cournot belief is the special case in which γ=0. The optimal γ* can be the value of γ that minimizes the distance[7] between stated beliefs[8] ($SB(t)$) and γ-weighted empirical beliefs ($B(t)$) in terms of mean square error.

Nyarko and Schotter (2002) investigated a belief-based learning model. They concluded that people behave in a manner consistent with belief learning.

---

7 $\min_\gamma \sum_{t-1}^{T} \left| SB(t) - B(t) \right|^2$

8 the true probability assigned to the chosen action until period t

Furthermore, they recommended that the types of belief used as inputs to these models need to be specified carefully. They found that the outcomes of stated beliefs differ from the outcomes of empirical beliefs, since unobserved factors affect the player's choice of behavior. This demonstrates how important it is for the model to capture the effect of unobserved interactions between competitors. By transforming an unobservable into an observable interaction, we can witness directly how parameter estimates change when new information is introduced (Nyarko and Schotter, 2002).

3.4.3. Reinforcement Models

Belief-based models assume that players hold beliefs about the likelihood of an opponent's action, and assume that the players choose their actions based upon their expected payoff, given these beliefs (Feltovich, 2000). On the other hand, reinforcement-based models do not require players to formulate any beliefs about their opponent's actions, since players do not have information about their opponents' type or actions. In reinforcement models, their strategies are reinforced relative to the payoffs they themselves earn over time, and players adjust their play to maximize these payoffs.

Mookherfee and Sopher (1997) and Erev and Roth (1998) compared the outcomes of a belief-based versus a reinforcement model. They found that a reinforcement model is better than a belief-based model in describing player behavior within certain types of game, even though enough information was given in the belief-based model used in their experiment. However, the generality of their results is questionable, since they only used data from their own experiments.

Sarin and Vahid (1999) conducted experiments to compare two reinforcement models. One is the *'fixed reference point'* model of Erev and Roth (1998), and the other their own *'SV'* model. The SV model assumes that a player chooses whatever strategy maximizes their payoffs, given the belief that the payoffs of each strategy are equal to the weighted averages of past payoffs. They concluded that an SV model can describe a player's behavior in a game better than fixed reinforcement models can.

Sarin and Vahid (1999) suggested a dynamic reinforcement learning model, which investigates how a player chooses action each time, given his or her own utility, which is updated with the player's own experience. Each time, players choose the strategy that yields the highest payoff. During this process, players ignore all future implications of their current choice, in terms of future choices and payoffs (Sarin and Vahid, 1999).

One form of SV reinforcement model is as follows (Sarin and Vahid, 1999):

$$U_n^j(t) = (1-\lambda)\cdot U_n^j(t-1) + \lambda \cdot \pi_n^j(\varpi) \quad (3.6)$$

$$U_n^k(t) = U_n^k(t-1) \text{ [9]}, \qquad \forall k \neq j \quad (3.7)$$

where

$U_n^j(t)$: reinforcement level of player n to choose strategy j at period t

λ: the proportional rate of player's surprise to previous reinforcement ($0<\lambda<1$)

[9] for the case of unchosen strategy at current time ($\lambda=0$)

$\pi_n^j(\varpi)$: player n 's payoff at period t from choosing strategy j given state of world (ϖ)

In this model, a player does not have information about the true choice environment, and does not update his or her subjective reinforcement about the payoff provided by other strategies. The above formulation shows that parameter λ determines how quickly the reinforcement approaches observed payoffs, and shows the effect of observed payoffs on reinforcement levels.

Sarin and Vahid (1999)'s SV model is similar to a belief-based model (fictitious play), since both models are assumed to be myopic, ignoring the implications of current choices on future choices and payoffs. One's beliefs are updated each period according to what one observes (Sarin and Vahid, 1999). However, the belief-based model assumes that beliefs are probabilistic, by observing the opponent's actions; whereas the SV model is concerned with the player's own payoffs from previous choices.

3.4.4. Mixed Models

Camerer and Ho (1998, 1999, 2002) developed a general experimental learning model (*experience-weighted attraction learning* (EWA) model) that incorporates elements from both the belief-based and reinforcement models. They illustrated that experimental learning models require the specification of initial attractions, how attractions are updated by experience, and how choice probabilities depend upon attractions. The EWA model assumes that each strategy has a numerical attraction, which determines the probability of a player choosing that strategy. In the experience-weighted attraction learning model, the attraction value, $A_i^j(t)$, is updated to be the sum of a depreciated experience-weighted previous attraction, $A_i^j(t-1)$, plus the weighted payoff from period t, normalized by the updated experience weight. Therefore, the attraction value and number of experiences are updated over time. The EWA model equation is as follows (Camerer and Ho, 2002):

$$A_n^j(t) = \frac{\phi \cdot N(t-1) \cdot A_n^j(t-1) + \left[\delta + (1-\delta) \cdot I(s_n^j, s_{-n}(t))\right] \cdot \pi_n(s_n^j, s_{-n}(t))}{N(t)} \tag{3.8}$$

$$N(t) = \rho \cdot N(t-1) + 1, \; t \geq 1 \tag{3.9}$$

where

n and $-n$: player

$S_n = \{s_n^1, s_n^2, ..., s_n^j\}$: discrete choice sets (strategy)

$A_n^j(t)$: attraction level of player n to choose strategy j at period t

$N(t)$: the number of 'observation-equivalents" of past experience (the unit of actual experience)

$N(0)$: the strength of initial attractions, relative to incremental changes in attractions due to actual experience and payoffs

$$I(s_n^j, s_{-n}^j) = \begin{cases} 1 & \text{if they choose } j \text{ strategy in period } t \\ 0 & \text{otherwise} \end{cases}$$

$\pi_n(s_n^j, s_{-n}(t))$: payoff function of player n at period t from choosing j strategy, which depends on the distribution of the opponents' price, denoted by $F_{-i}(p)$

ρ : depreciation rate that measures the fractional impact of previous experience, compared to one new period

$A_n^j(0)$: the initial attraction, which might be derived from similarity between strategies and strategies which were successful in similar games or from prior belief

δ : a discount factor for the payoff of unchosen choice

ϕ : a discount factor which depreciate previous attraction

The EWA model utilizes all the information provided by the players and the game environment, which clearly is different than the belief-based and reinforcement models. Camerer and Ho (1998, 1999, 2002) showed that belief learning and reinforc ement learning models are special cases of the EWA model. For this reason, they concluded that the EWA model combines both learning models and performs better than either of these two alternatives, estimating parameters using maximum likelihood estimations. Furthermore, the EWA model appeared to perform better than either pure model. However, the EWA model is incomplete, because it does not explain how a player's information about opponents' payoffs might influence decisions. In addition, this model does not capture the unobserved impact of competitive environment on bidding behavior and interaction among players' cognitive learning processes.

3.5. <u>Choice Models</u>

3.5.1. Multinomial Probit Models

Most previous choice theory models in the transportation field have utilized the multinomial logit (MNL) form. Two specific forms were used as the probabilistic choice rule in most of the previously-mentioned studies on learning: exponential (logit) and power. Although multinomial logit models have the extreme advantage of having a choice probability with a simple closed form, which can be calculated easily, the models nonetheless are limited by being *'independent of irrelevant alternatives'* (IIA). In order to overcome this problem, the *generalized extreme value model* (Train, 1986), the *heteroscedastic extreme value model* (Bhat, 1995), and the *multinomial probit model* (MNP) have been considered, and these models allow for a more flexible correlation structure of the error term, which is assumed to be normally distributed.

Among them, the MNP model provides the general framework to allow for interdependence of alternatives, with the most flexible pattern of error correlation structure in discrete choice analysis. Through this assumption, any error correlation can be postulated to capture the dynamic aspects of individual behavior, including state dependence, contemporaneous correlation, and taste variation (Jou, 1994). However, the MNP model is limited by computational difficulties associated with the evaluation of choice probabilities, including multidimensional normal integrals. The dynamic strategy choice model, under conditions of a competitive environment, also has multidimensional integrals for the multinomial density function, for which there is no closed form solution.

The dynamic multinomial probit model requires approximation methods to estimate parameters. Lerman and Manski (1981) first proposed a simulated maximum likelihood method by which to evaluate the multidimensional integrals required by the MNP model. A number of approximation methods have been proposed, but most approximation simulators have not been applied widely, because their results have not been accurate enough to satisfy empirical researchers.

Since the late 1980s, there have been steady advances in multinomial probit estimations (McFadden, 1989; Lam and Mahmassani, 1990; Bunch, 1991; Bolduc and Ben-Akiva, 1991; Bolduc, 1992; Geweke *et al,* 1994). McFadden (1989) proposed an efficient approximation method for multinomial probit applications, which exhibited computational efficiency in seeking model parameters. Lam and Mahmassani (1990) proposed a new MNP model estimation program, using a VMC (vectorized Monte Carlo) simulation procedure and new implementation of quasi-Newton BFGS (Broyden-Fletcher-Goldfarb-Shanno) nonlinear procedures to handle the large numbers of choice alternatives and general specifications. This estimation program has been applied successfully to dynamic travel behavior models with up to 17 alternatives (Mahmassani and Jou, 1996). Bunch (1991) simplified the multinomial probit model's covariance matrix; and Bolduc (1992) used auto-regressive errors for the multinomial probit estimation method with a large choice set, by means of simplifying its covariance matrix. Geweke *et al* (1994) proposed using a method of simulated moments, or alternatively using simulated maximum likelihood estimators with the GM recursive probability simulator, in order to estimate multinomial probit model parameters. Recently, Liu and Mahmassani (2000) presented a GAMNP model that incorporates genetic algorithms (GAs) and nonlinear programming (NLP) techniques to achieve a global optimum in maximum likelihood estimation.

The numerous improvements in the estimation algorithms and the simplification of the covariance matrix both encourage the application of multinomial probit models more broadly in the field, because they are fundamentally more flexible than the multinomial logit model. Therefore, the MNP model estimation program is applied to the estimation of dynamic competitive strategy choice model, and these estimation procedures are discussed further in Chapter 4.

3.5.2. Mixed Logit Models

Multinomial probit models have long been considered infeasible, because calculating the choice probabilities requires the evaluation of multiple integrals with no simple closed form solutions; in addition, their flexibility in the MNP model comes at a cost. Recent studies have shown that the mixed logit probability simulator can be one of the solutions to overcome problems in the MNP model, by means of an easy-to-compute and unbiased simulator. Adding the i.i.d. Gumbel term to the normal error terms leads to a particularly convenient and attractive probability simulator, which is the average of a set of logit probabilities. The mixed logit or kernel logit probability simulator has all of the desirable advantages of a simulator, which include being convenient, unbiased and smooth (Ben-Akiva, Bolduc, and Walker, 2001).

The mixed logit model was introduced both by Boyd and Mellman (1980) and by Cardell and Dunbar (1980). A more general model structure was required for the smooth probability simulators used in estimating mixed logit models. Since 1980, several papers have been published investigating various aspects of the mixed logit model, such that the model has become extremely popular in the literature (see McFadden 1989; Bolduc and Ben-Akiva 1991; Stern 1992; Bolduc, Fortin and

Fournier 1996; Bhat 1997 and 1998; Train 1998; Brownstone and Train 1999; Brownstone, Bunch and Train 2000; Goett, Hudson, and Train 2000; Srinivasan and Mahmassani 2000; Walker 2001 and 2004; and Srinivasan and Mahmassani 2005). It has been used in a wide variety of application areas.

Train (1998) and Bhat (1999) applied mixed logit formulation to capture heterogeneity in behavior across decision-makers for recreational demand and joint mode and departure time choice behaviors. These researchers demonstrated that the kernel logit model with normal errors is a more general form of mixed logit model, which can combine Gumbel errors with multivariate normal distributions (McFadden and Train 2000). This kernel logit model can be more flexible and realistic than the probit model, with fewer computational costs; it also allows for various model specifications.

While the number of logit kernel applications has been growing rapidly in the literature, the identification issue largely has been ignored. Recently, Ben-Akiva, Bolduc, and Walker (2001, 2004) presented a general framework for the specification, identification, and estimation of the logit kernel model. Srinivasan and Mahmassani (2003, 2005) also investigated the theoretical foundation of the dynamic kernel logit model.

A few investigators have applied the dynamic kernel logit model to longitudinal, discrete-choice data (Revelt and Train 1998; Goett, Hudson, and Train 2000). Srinivasan and Mahmassani (2005) utilized the dynamic kernel logit method to model longitudinal, discrete-choice data. They derived a dynamic kernel logit (DKL) formulation with normal errors for unordered discrete choice panel data. In addition, they presented the theoretical identification, suitability and properties of the kernel logit model, as well as the computational efficiency of DKL model performance. In this dissertation, the mixed logit model estimation procedure is applied for the dynamic strategy choice model under competitive environment. We compare those results to estimation results provided by the multinomial probit model in Chapters 6 and 7.

Chapter 4. Research Methodology

*4.1. **Introduction***

This dissertation is about the learning models used to describe a player's strategy choice behavior using experimental data. Our goal has been to explain, as accurately as possible and for every choice in an experiment, how that choice arises from the nature of multi-player interactions and their dynamics over multiple bids. Therefore, the principal focus of this dissertation is how to model a player's dynamic strategy choice behavior under the pressure of competition. We are most interested in which models describe human behavior best, when players make repeated decisions along with or in response to competitors' decisions. We propose a general dynamic strategy choice model framework and estimate the model parametrically, using two sets of experimental data.

This chapter presents the dynamic strategy choice modeling methodology to represent the dynamics of bidders' behavior and their learning processes in auction-based electronic freight marketplaces. As discussed previously, players in non-cooperative and competitive games try to assimilate a certain amount of feedback information concerning their opponents' bidding behavior, as they take part in similar auction games in a particular electronic market environment. This feedback information is an important factor in players' decision-making processes. Previous studies of dynamic choice behavior have not generally considered these interactions and associated dynamics among individuals in relation to the decision-making process. Introducing this concept of competition impact into dynamic choice model structures is this study's principal contribution to the development of a new perspective on transportation behavior modeling in competitive environments.

The modeling methodology refutes the generic assumptions of perfect player rationality and common knowledge. Different players may exhibit varying types or "degrees" of rationality, which depend on how much information feedback they receive, and their own experience. Between the two extremes of eductive and evolutionary learning mechanisms, this study explores two intermediate dynamic processes: the epistemic and behavioral reinforcement learning processes. As noted in the previous section, the epistemic process assumes that boundedly rational players revise incomplete beliefs about their opponents' behavior, while the behavioral reinforcement process assumes that players with bounded rationality choose flexible strategies in response to their own past results (Walliser, 1998). Two different types of dynamic strategy choice model structures are developed, one for each dynamic learning process considered.

The epistemic and behavioral reinforcement models have been treated as fundamentally different since 1950. Recently, a few researchers have asked whether the two learning models are related, based on the belief that people do not always apply the same learning process to every game. For this reason, two mixed learning models are formulated to explain the player's mixed learning behavior. The epistemic and behavioral reinforcement learning models are special cases of the mixed learning model that incorporate both kinds of information.

The dynamic multinomial probit (DMNP) and dynamic mixed logit (DML) models are applied in this dissertation to provide the mathematical framework for modeling bidders' dynamic strategy decisions. The DMNP and DML frameworks can be formulated for general discrete choice situations and can incorporate alternative behavioral theories, such as random utility maximization and bounded rationality. A key advantage of the DMNP and DKL model frameworks is that it allows flexible model specifications and realistic correlation structures for the analysis of dynamic discrete choices obtained from panel data. Of particular concern in this study is the effect of competition on multiple players' bidding behavior. Accordingly, new DMNP and DML model structures are developed to capture players' behavior in a competitive environment over time, including their reaction to the competitors' bidding strategy. DMNP and DML model framework for dynamic decisions in competitive environments is introduced next. The general structure of research methodology is illustrated in Figure 4.1.

The next section presents the general discrete choice model formulation used to compare the new model frameworks. In Section 4.3, the model specifications of epistemic learning and behavioral reinforcement learning are described. Section 4.4 shows the dynamic error structure used to represent the competitive interaction among players in their respective choice behaviors. The DMNP and DML model estimation procedures are discussed in Sections 4.5 and 4.6. Section 4.7 discusses the overall maximum likelihood estimation procedures for both models. In Section 4.8, the estimation procedure for mixed learning process is discussed.

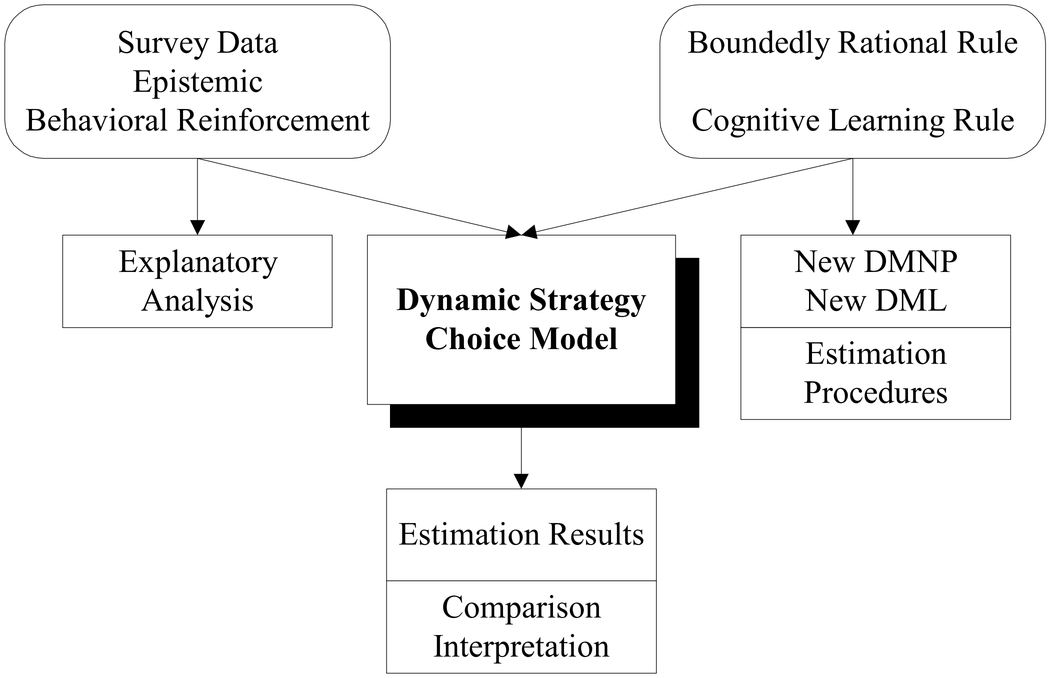


Figure 4.1 Overall structure of research methodology

4.2. ***General Discrete Choice Model Formulation***

Consider a player n, $n = 1,...,N$ where N is the number of players included in the sample, facing a set C_n of J_n discrete choice alternatives, denoted by i, $i = 1,...,J_n$; the model is written as:

$$d_n^{it} = \begin{cases} 1 & if \;\; U_n^{it} \geq U_n^{jt}, \; \forall j \neq i \\ 0 & \text{otherwise} \end{cases} \tag{4.1}$$

$$U_n^{it} = V_n^{it} + \varepsilon_n^{it} = \beta Z_n^{it} + \varepsilon_n^{it} \tag{4.2}$$

$$\varepsilon_n^{it} \sim MVN(0, \textstyle\sum_{\varepsilon})$$

where

U_n^{it} : the utility of alternative i as perceived by individual n

Z_n^{it} : a ($1\times K$) vector of explanatory variables for the utility to individual n of alternative i, including alternative specific dummy variables, as well as alternative specific attributes and individual characteristic variables

β : ($K\times 1$) vector of coefficients

ε_n^{it} : a random disturbance term (unobservable)

$E({\varepsilon_n^{it}}^2 = \sigma_i^2)$: variance of error term

$E(\varepsilon_n^{it}\varepsilon_n^{jt} = \rho_1)$: covariance across alternatives during the same time period

$E(\varepsilon_n^{it}\varepsilon_n^{it'} = \rho_2)$: covariance reflecting serial correlation over time periods

The assumption that the disturbances are multivariate normal distributed results in the probit and mixed logit model forms for the choice probability. With this specification, the general choice probability of alternative i for individual n is given by the following expression:

$$\begin{aligned} P_n^{it} &= prob(V_n^{it} + \varepsilon_n^{it} > V_n^{jt} + \varepsilon_n^{jt}), \forall j \neq i \\ &= \int d_n^{it}(V_n^{it} + \varepsilon_n^{it} > V_n^{jt} + \varepsilon_n^{jt}) f(\varepsilon_n) d\varepsilon_n, \forall j \neq i \end{aligned} \tag{4.3}$$

where

d_n^{it} : a binary indicator of whether the condition in parentheses holds.

This general DMNP and DML models are extended to develop the dynamic strategy choice model under a competitive environment.

4.3. Model Specification

This study develops the dynamic strategy choice model under a competitive environment, which captures players' behavior for both epistemic and behavioral reinforcement learning processes of four types of cognitive learning behavior introduced in Chapter 2. The model specifications for both cognitive learning models are described in this section and investigate how players' bidding reaction or behavior can differ according to the players' cognitive capacities. The utility in the epistemic learning model is decomposed into observed and unobserved parts:

$$\begin{aligned} U_n^{it} &= V_n^{it} + \beta_2 V_n^i(t-1) Y_n^{it} + \varepsilon_n^{it} \\ &= \beta_1 Y_n^{it} + \beta_0^i (1 - Y_n^{it}) \\ &+ \beta_1 \sum_{k=1}^{C_{-n}} \pi_n(s_n^i, s_{-n}^k(t)) \cdot p(s_{-n}^k(t) \mid s_{-n}^k(t-q), q = 1,\dots,t-1) + \beta_2 V_n^i(t-1)(1 - Y_n^{it}) + \varepsilon_n^{it} \end{aligned} \quad (4.4)$$

$$p(s_{-n}^k(t) \mid s_{-n}^k(t-q), q = 1,\dots,t-1) = \frac{N_{-n}^k(t-2) + I(s_{-n}^k(t-1))}{\sum_{h=1}^{C_{-n}} \left[N_{-n}^h(t-2) + I(s_{-n}^h(t-1)) \right]} \quad (4.5)$$

where

$\Im$: finite set of player = {1,2,…,n}

n and m : player, $n, m \in \Im$

$-n$: opponent player

h, i, j and k : alternative, $h, i, j, k \in C_n$

t : the time period when shipment r arrives and it is auctioned; auction epochs $T_r = \{t_1, t_2, \dots, t_q\}$

U_n^{it} : the utility of player n to choose alternative i at time (game) t

V_n^{it} : the deterministic term of the utility for player n at time (game) t for alternative i

ε_n^{it} : the normal error term component of the utility of alternative i at time (game) t to player n

β : ($K \times 1$) vector of coefficients on explanatory variables

β_I : the coefficient on initial utility at time (game) period t=0

β_0^i : the alternative specific constant for i

$$Y_n^{it} = \begin{cases} 1 & \text{if time period } t = 1 \text{ for alternative } i \\ 0 & \text{otherwise} \end{cases}$$

$s_n^i(t)$: the strategy of player n at time (game) t

$\pi_n(s_n^i, s_{-n}^k(t))$: a payoff when player n choose alternative i and opponent player $-n$ choose alternative k at time (game) period t

$I(s_n^i, s_n(t))$: the choice decision indicator:

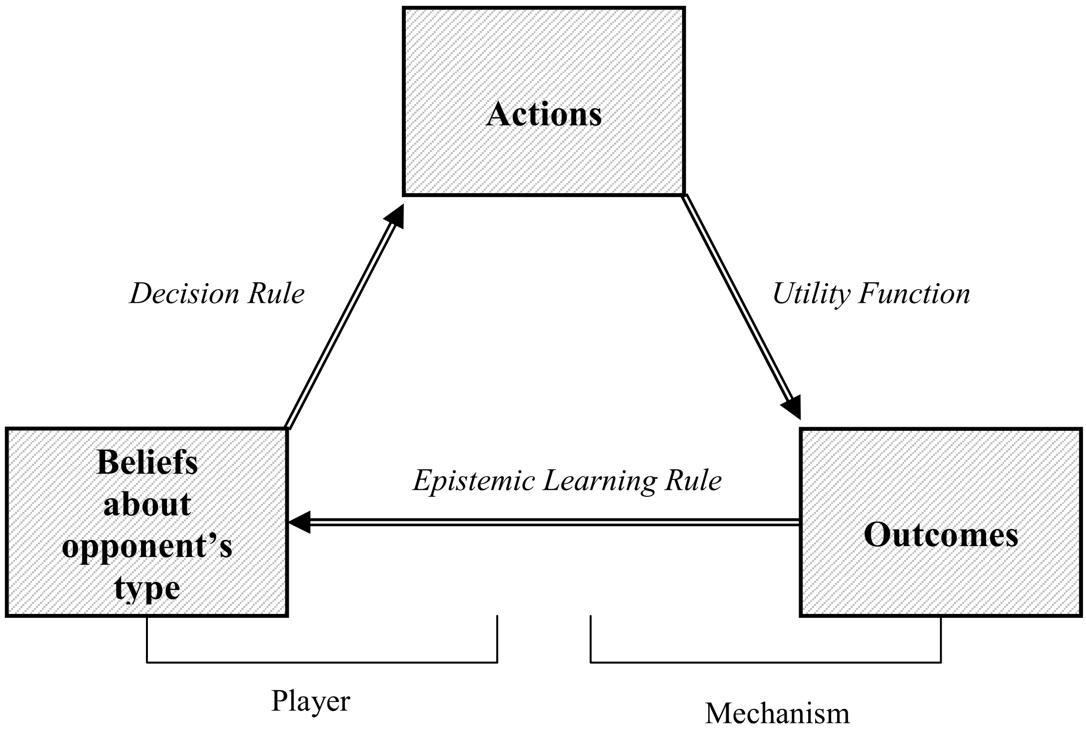
$$I(s_n^i, s_n(t-1)) = \begin{cases} 1 & \text{if player choose alternative } i \text{ at time period } t \text{ for } s_n^i = s_n^i(t-1) \\ 0 & \text{otherwise for } s_n^i \neq s_n^i(t-1) \end{cases}$$

$p(s_{-n}^k(t) | s_{-n}^k(t-q), q = 1,...,t-1)$: the probability of player n's belief about choices i of others $-n$ at time (game) period t

$N_n^i(t)$: the number of past auction participating experience about the choice i of player n at time (game) period t

If the player knows the other player's previous choice, his/her bidding decision can be affected by the player's belief regarding the opponent's choice. This model assumes that players would prefer a choice which has high expected payoffs, given beliefs formed by observing the history of what others did (Camerer and Ho, 1999). Players keep track of others' history from previous play, and form a function of player's belief about the opponent type on that game. Beliefs are updated by summing the previous number of experiences and one for the strategy combination the other players choose. The expected payoff value over time is achieved by multiplying that belief value for all strategies by payoffs at that time corresponding to a player's choice.

The following Figure 4.2 presents the conceptual framework for epistemic learning.



Source: Cheung and Friedman, 1997

Figure 4.2 Conceptual decision framework for epistemic learning

The other cognitive learning model considered in this study is the behavioral reinforcement learning model, which is written as:

$$
\begin{aligned}
U_n^{it} &= V_n^{it} + \beta_3 V_n^i(t-1)(1-Y_n^{it}) + \varepsilon_n^{it} \\
&= \beta_I Y_n^{it} + \beta_0^i(1-Y_n^{it}) + \beta_1 \left(\frac{\sum_{p=1}^{t-1} I(s_n^i(t-p)) \cdot \pi_n(s_n^i(t-p))}{\sum_{h=1}^{J_n} N_n^h(t-1)} \right) \\
&+ \beta_2 \sum_{p=1}^{t-1} \prod_{q=1}^{p} I(s_n^i(t-q)) + \beta_3 V_n^i(t-1)(1-Y_n^{it}) + \varepsilon_n^{it}
\end{aligned}
\tag{4.6}
$$

where

$$
I(s_n^i(t)) = \begin{cases} 1 & \text{if player choose alternative } i \text{ at time period } t \\ 0 & \text{otherwise} \end{cases}
$$

$\pi_n(s_n^i(t))$: a payoff when player n choose alternative i at time (game) period t

If a player does not have information about another player's history of decisions, he/she could only make a choice based on his/her experience. The reinforcement allows that player's choice decision to be directly reinforced by the previous results, and the propensity to opt for choice i depends on its stock of choice reinforcement (Arthur, 1991). The behavioral reinforcement learning model does not capture the direct impact of another player's decision on player own decision because he/she does not have records of another player's previous decisions over time. Based on the player's payoff over time periods, he/she can decide which alternative to choose.

The initial value (β_I) is included in both learning model specifications, and it reflects the influence of prior experience based upon a theory of first-period play and the player's own preference. This initial value can be estimated from the data. Our procedure is more general, because we estimate initial values as part of an overall maximization of utility.

Both epistemic and behavioral reinforcement model specifications include the internal habit persistence term, which is composed of two terms: 1) previous deterministic utility ($V_n^i(t-1)$); and 2) serial correlations (ρ_2) in the error term (see Section 4.4). *Habit persistence* is the relative weight given to the lagged utility in epistemic and behavioral reinforcement models. In this dissertation, the habit persistence term depends upon the history of past actions, and measures the sensitivity of the stock of habit to current action. In general, as the preference for future expected payoffs increases, the preference for habit persistence[10] decreases.

The conceptual decision framework for behavioral reinforcement learning is described in Figure 4.3.

---

[10] Braun, Constantidines and Ferson (1993) consider internal habit persistence specified in difference with one lag in action. More specifically, we assume that $U_n^i(t)$ takes the form: $U_n^i(t) = V_n^i(t) - \theta V_n^i(t-1)$ with $\theta \in (0,1)$

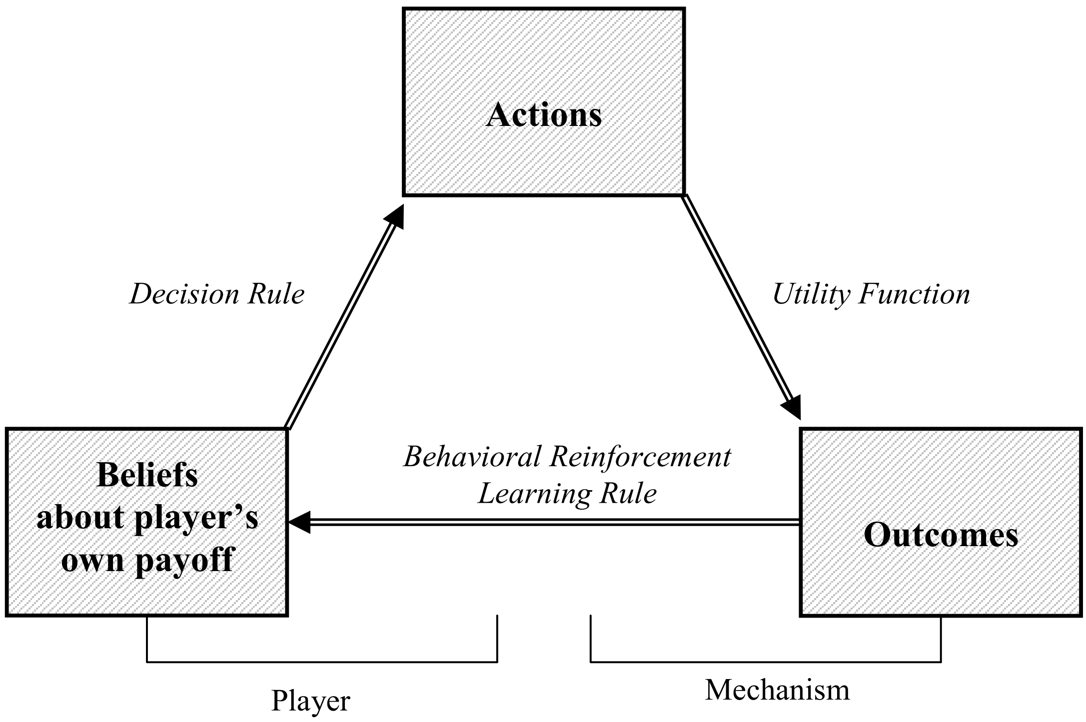


Figure 4.3 Conceptual decision framework for behavioral reinforcement learning

4.4. Dynamic Error Structure

In the model framework, the random utility term (ε_n^{it}) is made up of two components. The first component has been introduced in the probit disturbance term with a multivariate distribution, and it captures the interdependencies among the alternatives and shows the effect of serial correlation over time periods. The other error component captures the unobserved effect of players' competition with other players for winning the game. The disturbance term is specified as follows:

$$\varepsilon_n^{it} = v_n^{it} + \Omega_n^{it} \text{ for } n \neq m \tag{4.7}$$

$$v_n^{it} \sim MVN(0, \textstyle\sum_v)$$

$\Omega_n^{it} \sim MVN(0, \sum_{\gamma})$

where

$E(v_n^{it^2} = \sigma_i^2)$: the variance of error term

$E(v_n^{it} v_n^{jt} = \rho_1)$: the covariance across alternatives during the same time period

$E(v_n^{it} v_n^{it'} = \rho_2)$: the covariance reflecting serial correlation over time periods

$E(\Omega_n^{it} \Omega_m^{it} = \gamma_{nm}^{iit})$: the covariance across players for the same alternative

$E(\Omega_n^{it} \Omega_m^{jt} = \gamma_{nm}^{ijt})$: the covariance across players when player n choose alternative i and player m choose alternative j

The general MNP error term (v_n^{it}) captures unobserved preference heterogeneity across alternatives and serial correlation over time periods. The new error term (Ω_n^{it}) indicates the unobserved impact of player's decision behavior against the competitor's decision and it is specified as follows:

$$\gamma_{nm}^{ijt} \text{ for } n \neq m \tag{4.8}$$

Table 4.1 below shows the variance-covariance matrix in the case of two players, two alternatives, and two time (game) periods.

Table 4.1 Var-Cov Matrix for 2 players, 2 alternatives, and 2 time periods case

	Player	*m* =*1*			
Player	Alt time	*i* =*1* *t* =*1*	*i* =*2* *t* =*1*	*i* =*1* *t* =*2*	*i* =*2* *t* =*2*
n =*1*	*i* =*1* *t* =*1*	σ_1^2	ρ_1	ρ_2	0
	i =*2* *t* =*1*	ρ_1	σ_2^2	0	ρ_2
	i =*1* *t* =*2*	ρ_2	0	σ_1^2	ρ_1
	i =*2* *t* =*2*	0	ρ_2	ρ_1	σ_2^2

	Player	*m* =*2*			
Player	Alt time	*i* =*1* *t* =*1*	*i* =*2* *t* =*1*	*i* =*1* *t* =*2*	*i* =*2* *t* =*2*
n =*1*	*i* =*1* *t* =*1*	γ_{12}^{11}	γ_{12}^{12}	γ_{12}^{11}	γ_{12}^{12}
	i =*2* *t* =*1*	γ_{12}^{21}	γ_{12}^{22}	γ_{12}^{21}	γ_{12}^{22}
	i =*1* *t* =*2*	γ_{12}^{11}	γ_{12}^{12}	γ_{12}^{11}	γ_{12}^{12}
	i =*2* *t* =*2*	γ_{12}^{21}	γ_{12}^{22}	γ_{12}^{21}	γ_{12}^{22}

	Player	*m* =*1*			
Player	Alt time	*i* =*1* *t* =*1*	*i* =*2* *t* =*1*	*i* =*1* *t* =*2*	*i* =*2* *t* =*2*
n =*2*	*i* =*1* *t* =*1*	γ_{21}^{11}	γ_{21}^{12}	γ_{21}^{11}	γ_{21}^{12}
	i =*2* *t* =*1*	γ_{21}^{21}	γ_{21}^{22}	γ_{21}^{21}	γ_{21}^{22}
	i =*1* *t* =*2*	γ_{21}^{11}	γ_{21}^{12}	γ_{21}^{11}	γ_{21}^{12}
	i =*2* *t* =*2*	γ_{21}^{21}	γ_{21}^{22}	γ_{21}^{21}	γ_{21}^{22}

	Player	m =2			
Player	Alt time	i =1 t =1	i =2 t =1	i =1 t =2	i =2 t =2
n =2	i =1 t =1	σ_1^2	ρ_1	ρ_2	0
	i =2 t =1	ρ_1	σ_2^2	0	ρ_2
	i =1 t =2	ρ_2	0	σ_1^2	ρ_1
	i =2 t =2	0	ρ_2	ρ_1	σ_2^2

The summary of the error structure for dynamic strategy choice behavior is shown in Figure 4.4.

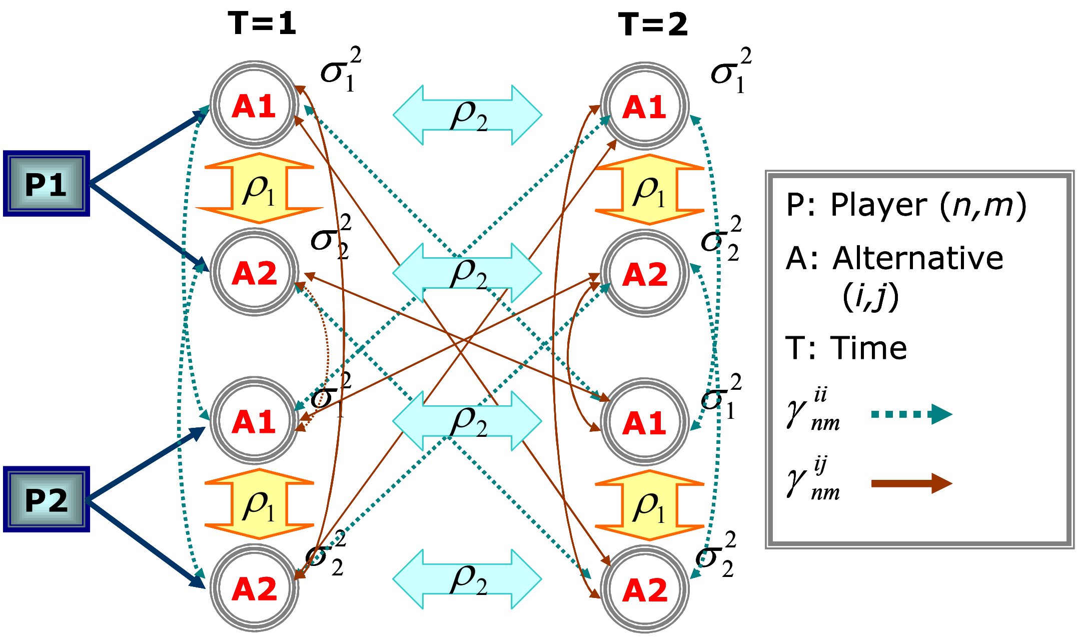


Figure 4.4 Summary of error structure for dynamic strategy choice behavior

4.5. Dynamic Multinomial Probit Model

Assuming that each player n chooses the set of four bid price alternatives i, DMNP model formulations for both dynamic learning models are based upon linear-in-parameter utilities, and are written as follows:

$$U_n^{it} = \beta Z_n^{it} + \varepsilon_n^{it} \tag{4.9}$$

with

$$d_n^{it} = \begin{cases} 1 & if \ U_n^{it} \geq U_n^{jt}, \ \forall j \neq i \\ 0 & \text{otherwise} \end{cases}$$

$$\varepsilon_n^{it} = v_n^{it} + \Omega_n^{it} \ \text{for} \ n \neq m \tag{4.10}$$

$$v_n^{it} \sim MVN(0, \textstyle\sum_v)$$

$$\Omega_n^{it} \sim MVN(0, \textstyle\sum_\gamma)$$

The random term (ε_n^{it}) is assumed to be jointly normally distributed, over time, and across different alternatives and individuals, with a zero mean and a general covariance matrix. ε_n^{it} is composed of $MT \times MT$ (where M = the number of alternative (bid prices) and T = the number of games (time periods) matrices that capture the correlations across alternatives, serial correlations due to the persistence of unobservable attributes across the sequence of games, and unobserved competitive influences on bidding behavior among individuals. One structure of the variance-covariance matrix in error term is as follows:

$E({v_n^{it}}^2 = \sigma_i^2)$: the variance of error term

$E(v_n^{it} v_n^{jt} = \rho_1)$: the covariance across alternatives during the same time period

$E(v_n^{it} v_n^{it'} = \rho_2)$: the covariance reflecting serial correlation over time periods

$E(\Omega_n^{it} \Omega_m^{it} = \gamma_{nm}^{iit})$: the covariance across players for the same alternative

$E(\Omega_n^{it} \Omega_m^{jt} = \gamma_{nm}^{ijt})$: the covariance across players when player n choose alternative i and player m choose alternative j

It can also be written as follow for the special case of Player 1's error structure with four alternatives and two time (game) periods:

Alternative Time	*i* =*1* *t* =*1*	*i* =*2* *t* =*1*	*i* =*3* *t* =*1*	*i* =*4* *t* =*1*
i =*1* *t* =*1*	$\sigma_1^2 + \gamma_{12}^{11}$	$\rho_1 + \gamma_{12}^{12}$	$\rho_1 + \gamma_{12}^{13}$	$\rho_1 + \gamma_{12}^{14}$
i =*2* *t* =*1*	$\rho_1 + \gamma_{12}^{12}$	$\sigma_2^2 + \gamma_{12}^{22}$	$\rho_1 + \gamma_{12}^{23}$	$\rho_1 + \gamma_{12}^{24}$
i =*3* *t* =*1*	$\rho_1 + \gamma_{12}^{13}$	$\rho_1 + \gamma_{12}^{23}$	$\sigma_2^2 + \gamma_{12}^{33}$	$\rho_1 + \gamma_{12}^{34}$
i =*4* *t* =*1*	$\rho_1 + \gamma_{12}^{14}$	$\rho_1 + \gamma_{12}^{24}$	$\rho_1 + \gamma_{12}^{34}$	$\sigma_2^2 + \gamma_{12}^{44}$

Alternative Time	*i* =*1* *t* =*2*	*i* =*2* *t* =*2*	*i* =*3* *t* =*2*	*i* =*4* *t* =*2*
i =*1* *t* =*1*	$\rho_2 + \gamma_{12}^{11}$	γ_{12}^{12}	γ_{12}^{13}	γ_{12}^{14}
i =*2* *t* =*1*	γ_{12}^{12}	$\rho_2 + \gamma_{12}^{22}$	γ_{12}^{23}	γ_{12}^{24}
i =*3* *t* =*1*	γ_{12}^{13}	γ_{12}^{23}	$\rho_2 + \gamma_{12}^{33}$	γ_{12}^{34}
i =*4* *t* =*1*	γ_{12}^{14}	γ_{12}^{24}	γ_{12}^{34}	$\rho_2 + \gamma_{12}^{44}$

In choice theory, the matter for choosing between two alternatives is the difference in utility between alternatives. Evaluating the log-likelihood function requires calculating the joint probability over all individuals in the sample associated with the chosen alternative. We need to derive the utilities into differences, with respect to the utility of the chosen alternative, so as to compute the joint choice probability simulator. Each time, we use the chosen alternative as the base. With this restriction, if alternative 3 is chosen, then the utility from alternative 3 is greater than the utility from alternatives 1, 2, and 4. Parameter β and the parameters in ε_n^{it} are not separately identified from observed choice behavior. Considering m alternatives to choose each game (time period), the new model can be written as:

$$U_n^{t^*} = \beta Z_n^{t^*} + \varepsilon_n^{t^*},\ \varepsilon_n^{t^*} \sim MVN(0, \textstyle\sum_\varepsilon) \tag{4.11}$$

$$U_n^{t^*} = \begin{bmatrix} U_n^{1t^*} \\ \dots \\ U_n^{mt^*} \end{bmatrix} = \begin{bmatrix} U_n^{1t} - U_n^{it} \\ \dots \\ U_n^{i-1,t} - U_n^{it} \\ U_n^{i+1,t} - U_n^{it} \\ U_n^{J,t} - U_n^{it} \end{bmatrix} \tag{4.12}$$

The general utility formulation is written in deviation with respect to the utility of the chosen alternative. The Cholesky decomposition formulation of the error term is derived to impose a positive definite error covariance matrix.

With the above error term and utility specifications, the joint choice probability ($\Pr\left(d_1^{it},....,d_n^{i't}\middle|\theta\right)$) of alternative i for individual n is given by the following DMNP formula:

$$\Pr\left(d_1^{it},....,d_n^{i't}\middle|\theta\right) = \Pr\left\{d_1^{it}(U_1^{jt} - U_1^{it}) \le 0,....,d_n^{i't}(U_n^{j't} - U_n^{i't}) \le 0\right\} \tag{4.13}$$

$(i, i', j, j' \in C_n,\ \forall i \ne j, and\ \forall i' \ne j')$

where

i and $i^{'}$: chosen alternative

j and $j^{'}$: unchosen alternative

$\theta = (\beta, E)$: the joint vector of parameters

β : the vector of unknown parameters in the systematic portion of the utility

E : the vector of unknown parameters in the error structure

$$d_n^{it} = \begin{cases} 1 & if \;\; U_n^{it} \geq U_n^{jt}, for\, i, j \in C_n, \;\; \forall i \neq j \\ 0 & \text{otherwise} \end{cases}$$

The difficulty with the DMNP model is that the resulting choice probabilities are multiple integrals. To solve this problem, the following DMNP estimation procedure is applied, using the Monte Carlo Simulation method.

First, make D draws of ε^d from the normal density; this process then is repeated D times. Second, calculate the joint choice probability ($\Pr\left(d_1^{it},....,d_n^{i^{'}t}\middle|\theta\right)$). The results of D draws are averaged to calculate the value of the log-likelihood function.

$$\hat{\Pr}\left(d_1^{it},....,d_n^{i^{'}t}\middle|\theta\right) = \frac{1}{D}\sum_{d=1}^{D}\Pr\left[d_1^{it},....,d_n^{i^{'}t}\middle|\left(\varepsilon_n^{it},\varepsilon_{-n}^{it}\right)^d\right] \tag{4.14}$$

The estimation method considered is based upon the maximization of the natural logarithm of the simulated likelihood function. The simulated log-likelihood function is written as:

$$\log L(\hat{\beta},\hat{E}) = \ln \hat{\Pr}\left(d_1^{it},....,d_n^{i^{'}t}\middle|\theta\right) \tag{4.15}$$

Once the simulated joint probabilities are obtained, the parameters in the above

DMNP model structure can be estimated using maximum likelihood. Those procedures are described in the Section 4.7 of this chapter.

4.6. Dynamic Mixed Logit Model

The multinomial probit (MNP) structure allows flexibility in the error structure for the covariance among the unobserved attributes of the alternatives. Unfortunately, in most choice contexts, the increase in flexibility of the MNP structure comes at the prohibitive cost of evaluating very high-dimensional multivariate normal integrals for choice probabilities. Over the past few years, researchers have discovered that simulation techniques, using a mixed or kernel logit model, can approximate the multi-dimensional integrals with smooth, unbiased and efficient simulators (Ben-Akiva *et al*, 2001).

The dynamic mixed logit model presented here is extended to consider mixed error structures, such as mixed or kernel logit variants of the MNP. This framework combines the flexibility and realism of the probit structure with some of the computational simplicity of the logit model. With the mixed logit framework, unobserved disturbance terms for each alternative can be divided into two components: a multivariate normally-distributed error component in the MNP, and a Gumbel-distributed error component in the MNL framework.

Gumbel error terms are assumed to be independent and identical over times, as well as across alternatives and individuals in the MNL model. Limitations in the i.i.d

Gumbel error term in MNL can be overcome by mixing the variance and covariance structure of the error term in MNP. Therefore, the Gumbel error terms in a mixed logit model can provide computational flexibility, by exploiting the closed logit-likelihood functional form, conditional upon the MVN error terms (Srinivasan and Mahmassani, 2005). The unconditional probability of choosing alternatives can be obtained by integrating the logit probability over the MVN error terms, by means of Monte Carlo integration. In this section, we discuss the calibration procedure for the dynamic mixed logit model. We develop the model formulation in the context of the model of a player's bid price choice. We use the same structures of utility specification and error structure for both cognitive learning processes as in the above DMNP model.

Assume M bid price choices, N individuals, and T time periods (games) in the choice set. Let the utility U_n^{it} that an individual associates with the bid price choice alternative be the sum of a deterministic component V_n^{it} that depends upon observed attributes of the alternative, as well as an individual and random component. In this discrete choice model, the utility that player n, $n = 1,..., N,$ where N is the sample size, receives from choosing alternative i, $i = 1, ..., J_n$ $(= M)$ is given by:

$$U_n^{it} = V_n^{it} + E_n^{it} = V_n^{it} + \varepsilon_n^{it} + \eta_n^{it} = V_n^{it} + v_n^{it} + \Omega_n^{it} + \eta_n^{it} \qquad (4.16)$$

The random vectors v_n^{it} and Ω_n^{it} are assumed to be mutually independent, and independent of η_n^{it}. The η_n^{it} terms are assumed to be independent and identically standard-Gumbel distributed across alternatives and individuals. In general, the dynamic mixed logit model accommodates a flexible covariance structure across alternatives and times. In our case, the dynamic mixed logit model can allow different correlations across individuals, which are expressed as follows:

$$E_n^{it} = \varepsilon_n^{it} + \eta_n^{it} \tag{4.17}$$

with

$\varepsilon_n^{it} = v_n^{it} + \Omega_n^{it}$ for $n \neq m$

$v_n^{it} \sim MVN(0, \sum_v)$

$\Omega_n^{it} \sim MVN(0, \sum_\gamma)$

$\eta_n^{it} \sim i.i.d.Gumbel(0, \sum_\eta)$

The random terms (ε_n^{it}) are assumed to be jointly, normally distributed, over time periods, alternatives and individuals, with a zero mean and a general covariance matrix. The ε_n^{it} matrices capture correlations across alternatives, serial correlations due to the persistence of unobservable attributes across the sequence of games, and unobserved competitive influences on bidding behavior among individuals, as in the above DMNP model. A structure of the variance-covariance matrix in the error term is as follows:

$E({v_n^{it}}^2 = \sigma_i^2)$: the variance of error term

$E(v_n^{it} v_n^{jt} = \rho_1)$: the covariance across alternatives during the same time period

$E(v_n^{it} v_n^{it'} = \rho_2)$: the covariance reflecting serial correlation over time periods

$E(\Omega_n^{it} \Omega_m^{it} = \gamma_{nm}^{iit})$: the covariance across players for the same alternative

$E(\Omega_n^{it}\Omega_m^{jt} = \gamma_{nm}^{ijt})$: the covariance across players when player n choose alternative i and player m choose alternative j

For given values of v_n^{it} and γ_n^{ijt} in the DML model, we get the familiar MNL form for the conditional probabilities that the choice i is given as ε_n^{it} and ε_{-n}^{it}:

$$P_n^t(i|\varepsilon_n^{it}, \varepsilon_{-n}^{it}) = \frac{\exp(V_n^{it} + \varepsilon_n^{it})}{\sum_{j \in C_n} \exp(V_n^{jt} + \varepsilon_n^{jt})} \quad \text{with } E_n^{it} = \varepsilon_n^{it} + \eta_n^{it} \tag{4.18}$$

Since ε_n^{it} is not known, the unconditional probability of choosing alternative i now can be obtained by integrating the conditional multinomial choice probabilities in equation (4.18) with respect to the assumed normal and independent distributions for the vectors v_n^{it} and γ_n^{ijt}:

$$P_n^t(i) = \int_{\varepsilon_n} p_n^t(i|\varepsilon_n^{it}, \varepsilon_{-n}^{it}) f(\varepsilon_n^{it}) d\varepsilon_n^{it} = \int_{\varepsilon_n} \left(\frac{\exp(V_n^{it} + \varepsilon_n^{it})}{\sum_{j \in C_n} \exp(V_n^{jt} + \varepsilon_n^{jt})} \right) f(\varepsilon_n^{it}) d\varepsilon_n^{it} \tag{4.19}$$

where $f(\varepsilon_n^{it})$ is the joint density function of ε_n^{it}, which is the sum of standard normals:

$$\varepsilon_n^{it} = v_n^{it} + \Omega_n^{it} \quad \text{for } n \neq m \tag{4.20}$$

For the maximum likelihood estimate, the probability of each sampled individual's sequence of observed choices over time is needed, which is:

$$S_n(i|\varepsilon_n^{it}, \varepsilon_{-n}^{it}) = \prod_t P_n^t(i|\varepsilon_n^{it}, \varepsilon_{-n}^{it}) = \prod_t \frac{\exp(V_n^{it} + \varepsilon_n^{it})}{\sum_{j \in C_n} \exp(V_n^{jt} + \varepsilon_n^{jt})} \tag{4.21}$$

The unconditional probability of the sequence of choices over time is written as:

$$S_n(i) = \int_{\varepsilon_n} S_n(i|\varepsilon_n^{it}, \varepsilon_{-n}^{it}) f(\varepsilon_n^{it}) d\varepsilon_n^{it} = \int_{\varepsilon_n} \prod_t P_n^t(i|\varepsilon_n^{it}, \varepsilon_{-n}^{it}) f(\varepsilon_n^{it}) d\varepsilon_n^{it} \qquad (4.22)$$

With the above error term and utility specifications, the joint choice probability among individuals, based upon the above probability of each sampled individual's sequence of observed choices over time of alternative i, is given by:

$$\Pr\left(d_1^{it}, \ldots, d_n^{i't} | \theta\right) = \Pr\left\{d_1^{it}(U_1^{jt} - U_1^{it}) \le 0, \ldots, d_n^{i't}(U_n^{j't} - U_n^{i't}) \le 0\right\} \qquad (4.23)$$

$$(i, i', j, j' \in C_n, \ \forall i \ne j, \ and \ \forall i' \ne j')$$

We rewrite this joint probability for DML model as:

$$\Pr\left(i|\varepsilon_n^{it}, \varepsilon_{-n}^{it}\right) = \prod_n \prod_t P_n^t(i|\varepsilon_n^{it}, \varepsilon_{-n}^{it}) = \prod_n \prod_t \frac{\exp(V_n^{it} + \varepsilon_n^{it})}{\sum_{j \in C_n} \exp(V_n^{jt} + \varepsilon_n^{jt})} \qquad (4.24)$$

where

i and i': chosen alternative

j and j': unchosen alternative

$\theta = (\beta, E)$: the joint vector of parameters

β: the vector of unknown parameters in the systematic portion of the utility

E: the vector of unknown parameters in the error structure (ε_n^{it})

$$d_n^{it} = \begin{cases} 1 & if \ U_n^{it} \ge U_n^{jt}, \ for \ i, j \in C_n, \ \forall i \ne j \\ 0 & \text{otherwise} \end{cases}$$

The unconditional probability of the sequence of choices over times is given by:

$$\Pr(i) = \int_{\varepsilon_n} \Pr(i|\varepsilon_n^{it}, \varepsilon_{-n}^{it}) f(\varepsilon_n^{it}) d\varepsilon_n^{it} \qquad (4.25)$$

We assume a linear-in-parameter specification for the systematic utility of each choice alternative, given by $V_n^{it} = \beta Z_n^{it}$ for individual n and alternative i. The composite random error term for individual n is given by $E_n^{it} = \varepsilon_n^{it} + \eta_n^{it} = v_n^{it} + \Omega_n^{it} + \eta_n^{it}$. We assume that v_n^{it}, Ω_n^{it} and η_n^{it} each are independently and identically distributed.

The purpose of estimation is to obtain model parameters by maximizing the likelihood function over all individuals in the sample. The parameters to be estimated in the DML model are the parameter vectors β in the systematic portion of utility, the variance-covariance matrix Σ_v and the variance-covariance matrix Σ_γ. The Cholesky decomposition formulation of the error term is derived to impose a positive definite error covariance matrix. Then, the log-likelihood function for a given value of the parameter vector ($\theta = (\beta, E)$) is written as:

$$L(\theta) = \ln \Pr(i) = \ln \int_{\varepsilon_n} \prod_n \prod_t \left(\frac{\exp(V_n^{it} + \varepsilon_n^{it})}{\sum_{j \in C_n} \exp(V_n^{jt} + \varepsilon_n^{jt})} \right) f(\varepsilon_n^{it}) d\varepsilon_n^{it} \tag{4.26}$$

Unfortunately, the log-likelihood function for the estimate of the parameters involves a multi-dimensional integral, which must be evaluated numerically, since it does not have a closed-form solution. The Monte Carlo simulation technique is applied to approximate the choice probabilities in the log-likelihood function of equation (4.26), by taking draws of $\theta = (\beta, E)$ from the population density and calculating the joint probability among individuals over time, and then maximizing the resulting simulated log-likelihood function. The average of $\Pr(i|\varepsilon_n^{it}, \varepsilon_{-n}^{it})$ over D draws yields the simulated probability among individuals:

$$\hat{\Pr}(i) = \frac{1}{D}\sum_{d=1}^{D} \Pr\left[i \middle| \left(\varepsilon_n^{it}, \varepsilon_{-n}^{it}\right)^d \right] \tag{4.27}$$

where

$\hat{\Pr}(i)$: the simulated joint choice probability among individuals choosing alternative i given parameter vector θ

$\left(\varepsilon_n^{it}, \varepsilon_{-n}^{it}\right)^d$: the dth realization of Monte-Carlo draws from the multi-variate normal distribution

This simulated joint probability is an unbiased estimator; and we expect that, as the number of draws (D) increases, the variance will decrease. This variance is smooth and sums to one over all alternatives (Train, 2003). The simulated log-likelihood of the sample is the sum of the natural logarithm of the simulated joint choice probabilities over all individuals:

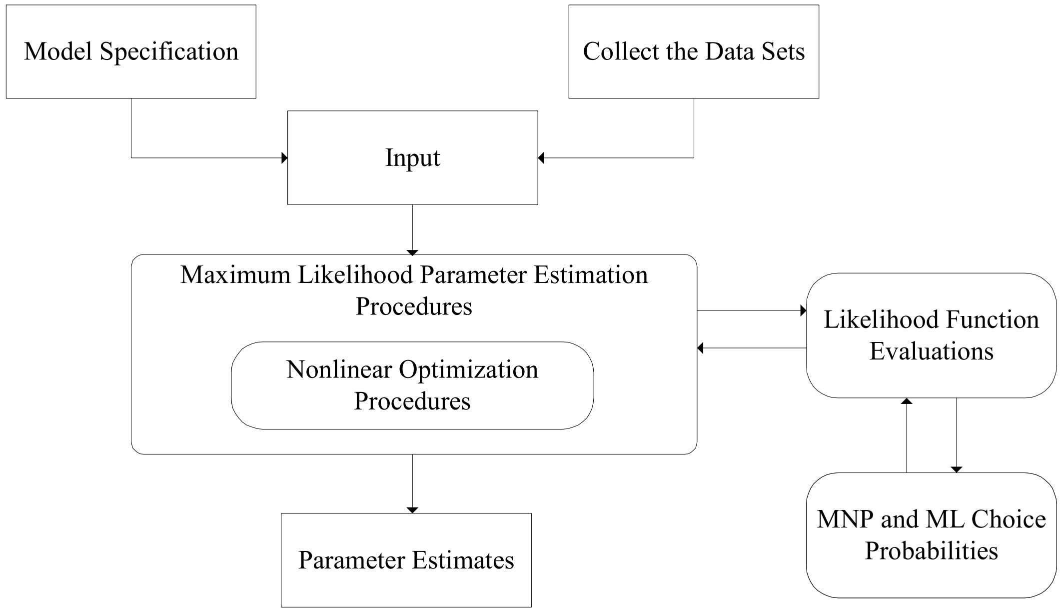
$$SL(\theta) = \ln \hat{\Pr}(i) = \ln\left\{ \frac{1}{D}\sum_{d=1}^{D} \Pr\left[i \middle| \left(\varepsilon_n^{it}, \varepsilon_{-n}^{it}\right)^d \right] \right\} \tag{4.28}$$

The above simulation technique approximates the joint choice probabilities. Since the elements within the vectors ε_n^{it}, including v_n^{it}, Ω_n^{it} and η_n^{it} are independent of each other, we generate a matrix ε_n^{it} of standard normal random numbers, and compute the corresponding joint choice probabilities for a given value of the parameter vector θ. The process is repeated D times for a given value of the parameter vector θ. The parameter vector θ is estimated as the vector value that maximizes the above-simulated function.

The maximum simulated log-likelihood (MSL) estimator has been considered to be simulated loglikelihood estimators, which is consistent, asymptotically efficient, and asymptotically normally distributed (Hajivassiliou, McFadden, and Ruud 1996; Lee 1992). The bias of the MSL estimator can be decreased, as the number of D draws increases. In the current study, we use 3000 repetitions for accurate simulations of the choice probabilities, and to reduce simulation variance of the MSL estimator.

4.7. Maximum Likelihood Estimation Process for Epistemic and Behavioral Reinforcement Learning Models

The general structure of the DMNP and DML models parameter estimation procedure is illustrated in Figure 4.5. The procedure shows the iterative approach to search for the maximum likelihood estimates. The likelihood function is evaluated each iteration, and the process is terminated when the specific convergence criterion is achieved (Jou and Mahmassani, 1994). The probit and mixed logit probability functions consist of multi-dimensional integrals that do not have closed-form solutions, and must be approximated numerically. Therefore, the Monte Carlo Simulation is applied to evaluate the both DMNP and DML choice probabilities. All estimations and computations are carried out using the FORTRAN programming language.



Source: Lam and Mahmassani (1991)

Figure 4.5 General Model Estimation Procedure

As mentioned earlier, the new DMNP model framework is presented to develop the dynamic strategy choice models for bidding behavior. To introduce competitive behaviors between players into the model estimation process, the following maximum likelihood parameter estimation process is described:

Step 1. Take the attribute sets and each individual's chosen alternative as inputs and specify the model structure.

Step 2. Specify the error structure to provide the error matrices for each individual.

1) Draw values of two error components from a normal density with a zero mean and covariance.

2) Apply the initial values for the parameters and the variance-

covariance terms.

3) Sum the two error components:

$\varepsilon_n^{it} = v_n^{it} + \Omega_n^{it}$

Step 3. 1 Reduce the dimension of the above variance-covariance matrices: $\varepsilon_n^{it^*} = (\varepsilon_n^{jt} - \varepsilon_n^{it})$

where i = the chosen alternative at time t, and j = the unchosen alternative at time t.

2) Provide the Cholesky decomposition matrix

Step 4. Use the configured error structure to generate realizations for the error components of the utilities.

Step 5. Calculate the deterministic utility values (V_n^{it})

Step 6. Reduce the dimension of the above deterministic utility: $V_n^{it^*} = (V_n^{jt} - V_n^{it})$

where, i : the chosen alternative at time t, and j : the unchosen alternative at time t.

Step 7. Monte Carlo Simulation

1) Using those values of errors and deterministic utility, calculate the total utility for each alternative: $U_n^{it^*} = V_n^{it^*} + \varepsilon_n^{it^*}$

2) Provide the joint choice probability, which is the likelihood of a sequence of decisions over all players and time periods:

$$\Pr\left(d_1^{it},\ldots.,d_n^{i't} \middle| \theta\right) = \Pr\left\{d_1^{it}(U_1^{jt} - U_1^{it}) \le 0,\ldots.,d_n^{i't}(U_n^{j't} - U_n^{i't}) \le 0\right\}$$

$(i, i^{'}, j, j^{'} \in C_n,\ \forall i \neq j, and\ \forall i^{'} \neq j^{'})$

3) Perform the Monte Carlo simulation until the choice probabilities of alternatives for each observation converge to steady values.

Step 8. Sum the probabilities of the chosen alternatives, which are the proportion of draws that are accepted:

$$\hat{\Pr}\left[d_1^{it},....,d_n^{i't}\middle|\theta\right] = \frac{1}{D}\sum_{d=1}^{D}\Pr\left[d_1^{it},....,d_n^{i't}\middle|\left(\varepsilon_n^{it},\varepsilon_{-n}^{it}\right)^d\right]$$

Step 9. Determine the next direction and step-size using the BFGS (Broyden-Fletcher-Goldfarb-Shanno) quasi-Newton algorithm with a central difference gradient to increase the non-linear likelihood function value (Jou and Mahmassani, 1994).

Step 10. Repeat the above steps until the convergence criterion is achieved.

In addition, the new DML model framework is explored to develop dynamic strategy choice models for the bidding behavior. The maximum likelihood parameter estimation procedure for the DML model is presented separately as follows:

Step 1~6. Same as in DMNP model estimation procedure

Step 7. Monte Carlo Simulation

1) Using the values of errors and deterministic utility, calculate the total utility for each alternative: $U_n^{it^*} = V_n^{it^*} + \varepsilon_n^{it^*}$

2) Provide the joint choice probability, which is the likelihood of a

sequence of decisions for all players over all time periods:

$$\Pr\left(i \middle| \varepsilon_n^{it}, \varepsilon_{-n}^{it}\right) = \prod_n \prod_t P_n^t\left(i \middle| \varepsilon_n^{it}, \varepsilon_{-n}^{it}\right) =$$

$$\prod_n \prod_t \left(\frac{1}{1 + \exp(U_n^{1t^*}) + \ldots + \exp(U_n^{(i-1)t^*}) + \exp(U_n^{(i+1)t^*}) + \ldots + \exp(U_n^{J_c t^*})} \right)$$

3) Perform Monte Carlo simulations until the choice probabilities of the alternatives for each observation converge to steady values.

Step 8. Sum the joint probabilities of the chosen alternatives, which is the proportion of draws that are accepted: $\hat{\Pr}(i) = \frac{1}{D}\sum_{d=1}^{D} \Pr\left[i \middle| \left(\varepsilon_n^{it}, \varepsilon_{-n}^{it}\right)^d \right]$

Step 9~10. Same as in DMNP model estimation procedure

The following Figure 4.6 shows the overall estimation procedure for the dynamic strategy choice model.

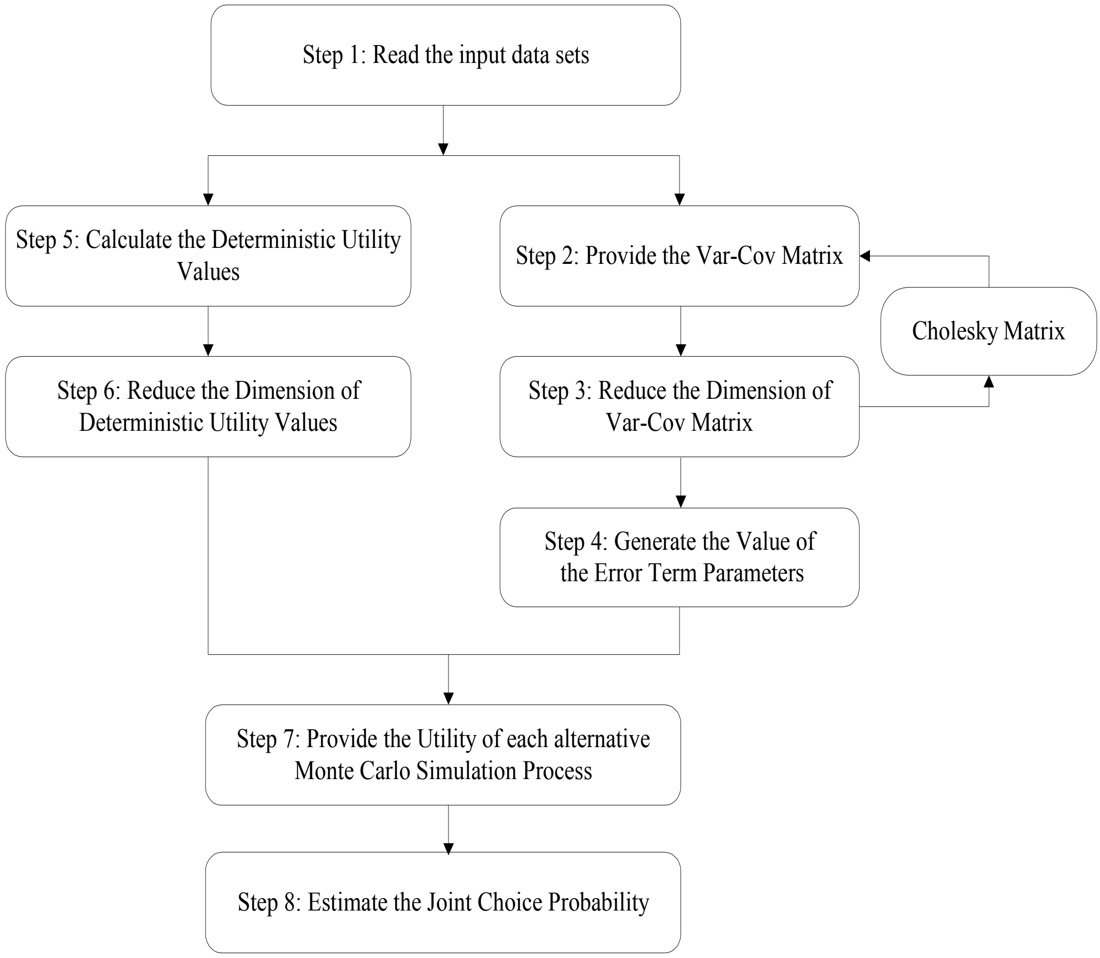


Figure 4.6 General Maximum Likelihood Parameter Estimation Procedures

4.8. Mixed Learning Model Specification and Estimation Process

We investigate two cognitive learning models in Section 4.1. These model specifications differ according to the level of cognitive capacity used for the given information. In this process, we expect that players make a decision based on their cognitive learning beliefs in sequential games. However, we question whether players continuously keep their learning beliefs for all games, if these players have enough information for epistemic learning game. People are not always rational when making decisions; hence, their decisions vary depending on the learning process for each game. Since players can change their beliefs or learning process based on their payoffs and their cognitive capacity, the player's strategy for each game can change with the corresponding either or both of cognitive learning process. This is called as the mixed learning process. Therefore, we propose the mixed learning model estimation procedure in this section and then estimate the model parametrically using epistemic learning datasets in Chapters 6 and 7.

The mixed learning model combines elements of two learning approach by including them as special cases. One approach is the epistemic learning model, which assumes that players keep track of the history of previous plays by others and form beliefs about what others will do based on past observation (Camerer and Anderson, 2000). Players make a best-response decision and expect that this strategy will maximize their expected payoff. In contrast, the behavioral reinforcement learning model assumes that strategies are reinforced by previous payoffs, and the propensity to choose a strategy depends on its reinforcement (Camerer and Anderson, 2000). Players care only about past payoffs, not about the mixed choice strategy of play that created those payoffs.

For a long time, researchers thought that these two approaches could not be combined since they rely on different cognitive capacity and information. Recently, however, some researchers have explored how the approaches might be related. The epistemic learning model does not consider the effect of past successes of chosen strategies, and the behavioral reinforcement model does not consider other players' reactions. Players make sequential decisions that rely on mixed beliefs created by their payoffs and the cognitive capacity to review other players' reactions. Based on this, the mixed learning model can combine appropriate elements of the epistemic and behavioral reinforcement learning approaches.

In the epistemic learning model, these mixed strategies, which represent the interdependence among players' choices, build players' beliefs about competitors' choices. The mixed learning model assumes that the player's decision is affected by competitors' learning beliefs, which decide the competitor's strategy. Therefore, mixed learning model structures are specified to provide each player's degree of propensity for following the mixed learning process (epistemic and behavioral reinforcement).

This dissertation proposes two mixed learning model structures. As previously mentioned, players do not always keep the same learning beliefs about the success or failure of a strategy; hence, players do not follow the same learning process in sequential games. In addition, players can change their decisions and learning process based on other players' reactions and their learning beliefs. Players can switch the strategic learning belief from either of the two learning processes to the other (epistemic and behavioral reinforcement). Therefore, two types of mixed learning models are formulated that have different assumptions about the interdependence of players' decisions with opponents' cognitive learning processes. One model makes an assumption about the interdependence among cognitive learning processes, while the other shows the independence of these processes.

We examine these mixed learning structures for parameters with clear psychological interpretations using statistical chi-square tests. To test the empirical usefulness of the mixed learning model, we derive maximum-likelihood parameter estimates from two epistemic learning datasets. The following process shows the first type of mixed learning model specification and estimation process. In this mixed learning model I, we consider that impact of independence among players' cognitive learning processes related to choice. In other words, the player's decision is not affected by other player's learning belief, which decides each competitor's decision. Therefore, we calculate each player's utility for each learning process and then provide the joint probability among players. We assume that players do not response to other competitor's cognitive learning belief in order to make a decision; thus, we do not consider the mixed strategy associated with the combination of players' cognitive learning processes to provide joint probability. We then verify the degree of impact of each learning process on each player's choice decision in the Monte Carlo simulation process. The utility specification for each learning model is same as in Section 4.1.

Step 1~4.	Same as in the above non-mixed learning model maximum likelihood estimation process
Step 5.	Calculate the deterministic utility values for each player

$$VE_n^{it} = \beta_I Y_n^{it} + \beta_0^i (1 - Y_n^{it})$$
$$+ \beta_1 \sum_{k=1}^{C_{-n}} \pi_n(s_n^i, s_{-n}^k(t)) \cdot p(s_{-n}^k(t) \big| s_{-n}^k(t-q), q = 1, \ldots, t-1) + \beta_4 VE_n^i(t-1)(1 - Y_n^{it}$$

$$VR_n^{it} = \beta_I Y_n^{it} + \beta_0^i (1 - Y_n^{it})$$
$$+ \beta_2 \left(\frac{\sum_{p=1}^{t-1} I(s_n^i (t-p)) \cdot \pi_n (s_n^i (t-p))}{\sum_{h=1}^{J_n} N_n^h (t-1)} \right) + \beta_3 \sum_{p=1}^{t-1} \prod_{q=1}^{p} I(s_n^i (t-q))$$
$$+ \beta_4 VR_n^i (t-1)(1 - Y_n^{it})$$

where

VE_n^{it} : player n 's deterministic utility with epistemic learning rule of choice i at game (time) t

VR_n^{it} : player n 's deterministic utility with behavioral reinforcement learning rule of choice i at game (time) t

Step 6. Reduce the dimension of the above player n 's deterministic utility for each learning process

$$VE_n^{it^*} = (VE_n^{jt} - VE_n^{it})$$

$$VR_n^{it^*} = (VR_n^{jt} - VR_n^{it})$$

where, i : the chosen alternative at time t , and j : the unchosen alternative at time t .

Step 7. Monte Carlo Simulation

1) Using those values of errors and deterministic utility, calculate the total utility for each alternative and each draw:

$$UE_n^{D_{En}it^*} = VE_n^{D_{En}it^*} + \varepsilon_n^{D_{En}it^*},\ UR_n^{D_{Rn}it^*} = VR_n^{D_{Rn}it^*} + \varepsilon_n^{D_{Rn}it^*}$$

where

$UE_n^{D_{En}it^*}$: player n 's difference in utility between chosen i and unchosen j for epistemic learning process at game (time) t for D_{En} th draw

$UR_n^{D_{Rn}it^*}$: player n 's difference in utility between chosen i and unchosen j for behavioral reinforcement learning process at game (time) t for D_{Rn} th draw

2) Provide the joint choice probability of each learning process among players, which is the likelihood of a sequence of decisions over all players and time periods.

two-player experiment:

$$P\left(d_1^{it}, d_2^{i't} \middle| \theta\right) = \Pr\left\{d_1^{it}(U_1^{jt} - U_1^{it}) \le 0, d_2^{i't}(U_2^{j't} - U_2^{i't}) \le 0\right\}$$

three-player experiment:

$$P\left(d_1^{it}, d_2^{i't}, d_3^{i''t} \middle| \theta\right) = \Pr\left\{\begin{matrix} d_1^{it}(U_1^{jt} - U_1^{it}) \le 0, d_2^{i't}(U_2^{j't} - U_2^{i't}) \le 0, \\ d_3^{i''t}(U_1^{j''t} - U_1^{i''t}) \le 0 \end{matrix}\right\}$$

$(i, i', i'', j, j', j'' \in C_n,\ \forall i \ne j,\ \forall i' \ne j', and\ \forall i'' \ne j'')$

where

U_n^{it} : player n 's utility with either of epistemic or behavioral reinforcement learning rule of choice i at game (time) t

i , i' and i'': chosen alternative

j , j' and j'': unchosen alternative

$\theta = (\beta, E)$: the joint vector of parameters

3) Perform the Monte Carlo simulation until the choice probabilities of alternatives for each observation converge to steady values.

Step 8. Sum the probabilities of each learning process of the chosen alternatives, which are the proportion of draws that are accepted:

$$\hat{\Pr}\left(d_1^{it},\ldots,d_n^{i't}\middle|\theta\right)=\frac{1}{D}\sum_{d=1}^{D}\Pr\left[d_1^{it},\ldots,d_n^{i't}\middle|\left(\varepsilon_n^{it},\varepsilon_{-n}^{it}\right)^d\right]$$

Two-player experiment:

$$\hat{P}\left[d_1^{it},d_2^{i't}\middle|\theta\right]=\frac{1}{D}\left\{\sum_{d=1}^{D}P\left[\begin{pmatrix} d_1^{it}(UE_1^{D_1^E jt}-UE_1^{D_1^E it})\le 0,\\ d_1^{it}(UR_1^{D_1^R jt}-UR_1^{D_1^R it})\le 0,\\ d_2^{i't}(UE_2^{D_2^E j't}-UE_2^{D_2^E i't})\le 0,\\ d_2^{i't}(UR_2^{D_2^R j't}-UR_2^{D_2^R i't})\le 0 \end{pmatrix}\middle|\left(\varepsilon_n^{it},\varepsilon_{-n}^{it}\right)^d\right]\right\}$$

Three-player experiment:

$$\hat{P}\left[d_1^{it},d_2^{i't},d_3^{i''t}\middle|\theta\right]=\frac{1}{D}\left\{\sum_{d=1}^{D}P\left[\begin{pmatrix} d_1^{it}(UE_1^{D_1^E jt}-UE_1^{D_1^E it})\le 0,\\ d_1^{it}(UR_1^{D_1^R jt}-UR_1^{D_1^R it})\le 0,\\ d_2^{i't}(UE_2^{D_2^E j't}-UE_2^{D_2^E i't})\le 0,\\ d_2^{i't}(UR_2^{D_2^R j't}-UR_2^{D_2^R i't})\le 0,\\ d_3^{i't}(UE_3^{D_3^E j''t}-UE_3^{D_3^E i''t})\le 0,\\ d_3^{i't}(UR_3^{D_3^R j''t}-UR_3^{D_3^R i''t})\le 0 \end{pmatrix}\middle|\left(\varepsilon_n^{it},\varepsilon_{-n}^{it}\right)^d\right]\right\}$$

$D_1^E=\alpha_1^E D$ and $D_1^R=D-D_1^E=(1-\alpha_1^E)D$ for Player 1

$D_2^E=\alpha_2^E D$ and $D_2^R=D-D_2^E=(1-\alpha_2^E)D$ for Player 2

$D_3^E=\alpha_3^E D$ and $D_3^R=D-D_3^E=(1-\alpha_3^E)D$ for Player 3

where

D_n^E : player n's number of draws for epistemic learning rule

D_n^R : player n's number of draws for behavioral reinforcement learning rule

D : total number of Monte Carlo simulation draws

α_n^E : parameter value for player n 's degree of propensity for the epistemic learning process over all games

$1-\alpha_n^E$: parameter value for player n 's degree of propensity for the behavioral reinforcement learning process over all games

Step 9~10. Same as in the above non mixed learning model maximum likelihood estimation process

The next process presents the second type of mixed learning model specification and estimation process. The mixed learning model II considers the impact of interdependence among players' learning rules related to choice decision. Players can switch their learning processes or choices according to competitors' learning beliefs, which influence their strategies. Based on this hypothesis, we provide each player's utility and the joint probability for the mixed combination of players' learning processes. In this process, the interdependence impact of the mixed combination of learning processes among players can be investigated by providing the sum of the partial joint probability for each mixed learning process. The following procedure shows the mixed learning model II estimation process.

Step 1~6. Same as in the above mixed learning model I estimation process

Step 7. Monte Carlo Simulation

1) Using those values of errors and deterministic utility, calculate the

total utility for each alternative:

$UE_n^{it^*} = VE_n^{it^*} + \varepsilon_n^{it^*}$, $UR_n^{it^*} = VR_n^{it^*} + \varepsilon_n^{it^*}$

where

$UE_n^{it^*}$: player n's difference in utility between chosen i and unchosen j for epistemic learning process at game (time) t for D_{En} th draw

$UR_n^{it^*}$: player n's difference in utility between chosen i and unchosen j for behavioral reinforcement learning rule of choice i at game (time) t

2) Provide the partial joint choice probability of each combination of learning process among players, which is the likelihood of a sequence of decisions over all players and time (game) periods.

Two-player experiment:

$$PEE\left(d_1^{it}, d_2^{i't} \middle| \theta\right) = \Pr\left\{d_1^{it}(UE_1^{jt} - UE_1^{it}) \leq 0, d_2^{i't}(UE_2^{j't} - UE_2^{i't}) \leq 0\right\}$$

$$PER\left(d_1^{it}, d_2^{i't} \middle| \theta\right) = \Pr\left\{d_1^{it}(UE_1^{jt} - UE_1^{it}) \leq 0, d_2^{i't}(UR_2^{j't} - UR_2^{i't}) \leq 0\right\}$$

$$PRE\left(d_1^{it}, d_2^{i't} \middle| \theta\right) = \Pr\left\{d_1^{it}(UR_1^{jt} - UR_1^{it}) \leq 0, d_2^{i't}(UE_2^{j't} - UE_2^{i't}) \leq 0\right\}$$

$$PRR\left(d_1^{it}, d_2^{i't} \middle| \theta\right) = \Pr\left\{d_1^{it}(UR_1^{jt} - UR_1^{it}) \leq 0, d_2^{i't}(UR_2^{j't} - UR_2^{i't}) \leq 0\right\}$$

Three-player experiment:

$$PEEE\left(d_1^{it}, d_2^{i't}, d_3^{i''t} \middle| \theta\right) = \Pr\left\{\begin{matrix} d_1^{it}(UE_1^{jt} - UE_1^{it}) \leq 0, d_2^{i't}(UE_2^{j't} - UE_2^{i't}) \leq 0, \\ d_3^{i''t}(UE_3^{j''t} - UE_3^{i''t}) \leq 0 \end{matrix}\right\}$$

$$PEER\left(d_1^{it}, d_2^{i't}, d_3^{i''t} \middle| \theta\right) = \Pr\left\{\begin{matrix} d_1^{it}(UE_1^{jt} - UE_1^{it}) \le 0, d_2^{i't}(UE_2^{j't} - UE_2^{i't}) \le 0, \\ d_3^{i''t}(UR_3^{j''t} - UR_3^{i''t}) \le 0 \end{matrix}\right\}$$

$$PERR\left(d_1^{it}, d_2^{i't}, d_3^{i''t} \middle| \theta\right) = \Pr\left\{\begin{matrix} d_1^{it}(UE_1^{jt} - UE_1^{it}) \le 0, d_2^{i't}(UR_2^{j't} - UR_2^{i't}) \le 0, \\ d_3^{i''t}(UR_3^{j''t} - UR_3^{i''t}) \le 0 \end{matrix}\right\}$$

$$PERE\left(d_1^{it}, d_2^{i't}, d_3^{i''t} \middle| \theta\right) = \Pr\left\{\begin{matrix} d_1^{it}(UE_1^{jt} - UE_1^{it}) \le 0, d_2^{i't}(UR_2^{j't} - UR_2^{i't}) \le 0, \\ d_3^{i''t}(UE_3^{j''t} - UE_3^{i''t}) \le 0 \end{matrix}\right\}$$

$$PREE\left(d_1^{it}, d_2^{i't}, d_3^{i''t} \middle| \theta\right) = \Pr\left\{\begin{matrix} d_1^{it}(UR_1^{jt} - UR_1^{it}) \le 0, d_2^{i't}(UE_2^{j't} - UE_2^{i't}) \le 0, \\ d_3^{i''t}(UE_3^{j''t} - UE_3^{i''t}) \le 0 \end{matrix}\right\}$$

$$PRRE\left(d_1^{it}, d_2^{i't}, d_3^{i''t} \middle| \theta\right) = \Pr\left\{\begin{matrix} d_1^{it}(UR_1^{jt} - UR_1^{it}) \le 0, d_2^{i't}(UR_2^{j't} - UR_2^{i't}) \le 0, \\ d_3^{i''t}(UE_3^{j''t} - UE_3^{i''t}) \le 0 \end{matrix}\right\}$$

$$PRER\left(d_1^{it}, d_2^{i't}, d_3^{i''t} \middle| \theta\right) = \Pr\left\{\begin{matrix} d_1^{it}(UR_1^{jt} - UR_1^{it}) \le 0, d_2^{i't}(UE_2^{j't} - UE_2^{i't}) \le 0, \\ d_3^{i''t}(UR_3^{j''t} - UR_3^{i''t}) \le 0 \end{matrix}\right\}$$

$$PRRR\left(d_1^{it}, d_2^{i't}, d_3^{i''t} \middle| \theta\right) = \Pr\left\{\begin{matrix} d_1^{it}(UR_1^{jt} - UR_1^{it}) \le 0, d_2^{i't}(UR_2^{j't} - UR_2^{i't}) \le 0, \\ d_3^{i''t}(UR_3^{j''t} - UR_3^{i''t}) \le 0 \end{matrix}\right\}$$

3) Provide the total joint probability for all combinations of learning processes among players

$$P\left(d_1^{it}, d_2^{i't} \middle| \theta\right) = \alpha_{12}^{EE} PEE\left(d_1^{it}, d_2^{i't} \middle| \theta\right) + \alpha_{12}^{ER} PER\left(d_1^{it}, d_2^{i't} \middle| \theta\right)$$
$$+ \alpha_{12}^{RE} PRE\left(d_1^{it}, d_2^{i't} \middle| \theta\right) + \alpha_{12}^{RR} PRR\left(d_1^{it}, d_2^{i't} \middle| \theta\right)$$

for two-player

$$P\left(d_1^{it}, d_2^{i't}, d_3^{i''t} \middle| \theta\right) = \alpha_{123}^{EEE} PEEE\left(d_1^{it}, d_2^{i't}, d_3^{i''t} \middle| \theta\right) + \alpha_{123}^{EER} PEER\left(d_1^{it}, d_2^{i't}, d_3^{i''t} \middle| \theta\right) +$$
$$\alpha_{123}^{ERE} PERE\left(d_1^{it}, d_2^{i't}, d_3^{i''t} \middle| \theta\right) + \alpha_{123}^{ERR} PERR\left(d_1^{it}, d_2^{i't}, d_3^{i''t} \middle| \theta\right) +$$
$$\alpha_{123}^{REE} PREE\left(d_1^{it}, d_2^{i't}, d_3^{i''t} \middle| \theta\right) + \alpha_{123}^{RER} PRER\left(d_1^{it}, d_2^{i't}, d_3^{i''t} \middle| \theta\right) +$$
$$\alpha_{123}^{RRE} PRRE\left(d_1^{it}, d_2^{i't}, d_3^{i''t} \middle| \theta\right) + \alpha_{123}^{RRR} PRRR\left(d_1^{it}, d_2^{i't}, d_3^{i''t} \middle| \theta\right)$$

for three-player

$(i, i', i'', j, j', j'' \in C_n,\ \forall i \neq j,\ \forall i' \neq j', and\ \forall i'' \neq j'')$

$$\alpha_{12}^{EE} + \alpha_{12}^{ER} + \alpha_{12}^{RE} + \alpha_{12}^{RR} = 1$$

$$\alpha_{123}^{EEE} + \alpha_{123}^{EER} + \alpha_{123}^{ERE} + \alpha_{123}^{ERR} + \alpha_{123}^{REE} + \alpha_{123}^{RER} + \alpha_{123}^{RRE} + \alpha_{123}^{RRR} = 1$$

where

i , i' and i'': chosen alternative

j , j' and j'': unchosen alternative

$\theta = (\beta, E)$: the joint vector of parameters

α_n^m : parameter value for player's degree of propensity for the m type of mixed cognitive learning process over all games

3) Perform the Monte Carlo simulation until the choice probabilities of alternatives for each observation converge to steady values.

Step 8. Sum the probabilities of the chosen alternatives, which are the proportion of draws that are accepted:

$$\hat{P}\left[d_1^{it}, \ldots, d_n^{i't} \middle| \theta\right] = \frac{1}{D} \sum_{d=1}^{D} P\left[d_1^{it}, \ldots, d_n^{i't} \middle| \left(\varepsilon_n^{it}, \varepsilon_{-n}^{it}\right)^d\right]$$

Step 9~10. Same as in the above mixed learning model I estimation process

Chapter 5. Explanatory Analysis

5.1. Introduction

The previous chapter discussed the research methodology of the current study, including the epistemic and behavioral learning modeling frameworks using the dynamic multinomial probit (DMNP) and dynamic mixed logit (DML) structure. Those model structures were applied to data sets obtained from two experiments in which decision-makers are in hypothetical bidding situations. The first experiments included two players who separately participated in the two types of games (epistemic and behavioral reinforcement) 80 games (times) each, and who were able to choose between four discrete alternative bid prices each time. The former type of game provided subjects with enough information for behavioral reinforcement, but not enough for epistemic learning. The latter game provided subjects with enough information for epistemic learning. The first price auction[11] was applied to the first experiment introduced.

To extend our modeling structure and evaluate the performance of the dynamic strategy choice model, the second experiment is necessary. In these experiments, three people play a sequential auction game 80 times with the same opponent under

---

[11] In a procurement auction, the winner is the bidder who submits the lowest bid, and is paid an amount equal to his or her bid. (source: http://www.gametheory.net/Dictionary/)

various treatments/scenarios. The second price auction[12] was applied. The estimation results for DML model is provided using the data from this second experiment. Those results were compared to the estimation results using the first experiment data.

Each player was asked to provide a choice decision among the four bid prices for the two types of game. For the behavioral reinforcement learning case, players did not have information about the opponent's previous choice; they only knew their own history of decisions, and who had won each prior game. In the epistemic learning survey, each player was notified after each game about all players' choices. Players were able to track their opponent's decision history as well their own decisions.

The rule for winning was that whichever player bid the lower bid price would win the game. Each player sustained different costs during each game and his or her payoff depended upon the different costs each time, since the payoff was the difference between the chosen bid price and the cost at any given time. If both players bid the same price, a given player's payoff is half the difference between the chosen bid price and the corresponding cost. The player's payoff and history of choice decision index were used as alternative specific explanatory variables. Estimations were performed for epistemic, behavioral reinforcement, and mixed learning models.

Note that while the decision-makers are professionals, they do not work for freight

---

[12] In a procurement auction, the winner is the bidder who submits the lowest bid, and is paid an amount equal to the next lowest submitted bid. (source: http://www.gametheory.net/Dictionary/)

carriers, nor are their responses intended for any other purpose than to illustrate application of the models introduced in this study, and how the parameter values would be interpreted in light of the particular specifications of interest.

In this study, it is assumed that private and public information are the same across different games. A player's cost information only is revealed to the player him or herself. Accordingly, the payoff also is private information.

5.2. Input Data Description

As previously mentioned, the datasets were collected through two experiments designed to examine epistemic and behavioral reinforcement learning. The first experiment included two players in a first price auction, while the second experiment consisted of three players in a second price auction. These data were derived from experiments involving 80 repetitions of two games under different informational assumptions and payoff structures.

All games were designed to allow players to choose from four bid prices with different assumptions for each of two players. In each game, each player's bid price and cost determine that player's payoff. Table 5.1 shows the bid price index as a choice alternative. We used a uniform distribution to generate random cost values between 10 and 20 for each player. Each player's costs are different each game, as is shown in Figures 5.1 and 5.2.

Table 5.1 Choice set for bid price

Index	Bid Price (BP)
Alternative 1	10
Alternative 2	15
Alternative 3	20
Alternative 4	25

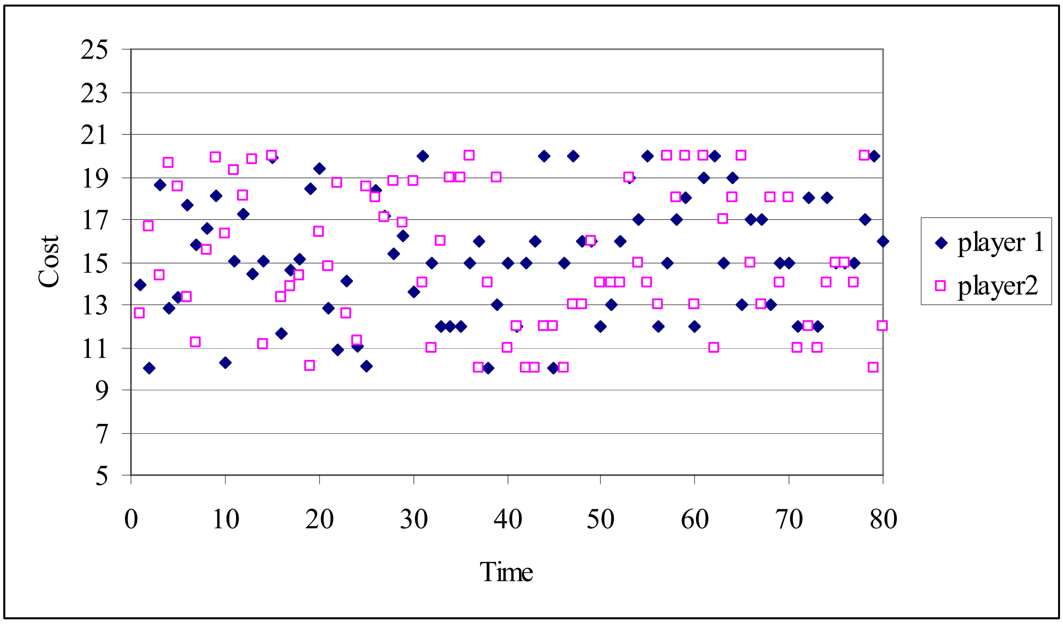


Figure 5.1 Players' costs for each time period (two-player experiment)

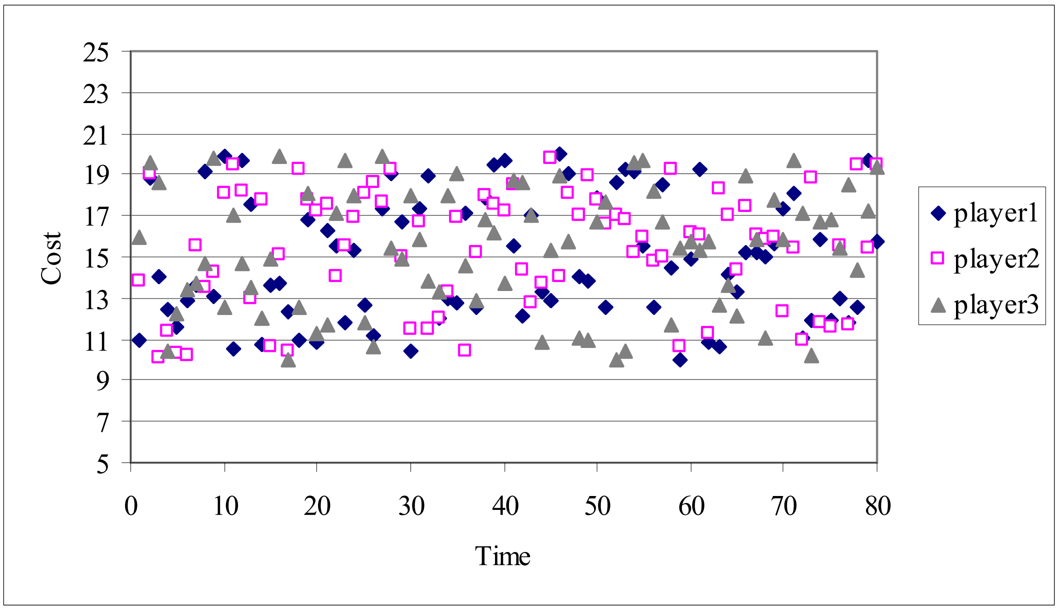


Figure 5.2 Players' costs for each time period (three-player experiment)

Payoff in these games (for players) is a critical factor to determine the player's bidding strategy. The formulation to calculate payoffs for the first experiment is:

$$\pi_n^{it} = \begin{cases} BP_n^{it} - Cost_n^t & \text{If each player chooses the different bid price alternative} \\ \left(\dfrac{BP_n^{it} - Cost_n^t}{2} \right) & \text{Otherwise} \end{cases} \tag{5.1}$$

where

π_n^{it} : player n's payoff for choosing the alternative bid price i at time (game) t

BP_n^{it} : player n's bid price as alternative i at time (game) t

$Cost_n^t$: player n's cost at time (game) t

The following tables, Tables 5.2 and 5.3, show each player's payoffs using the mixed strategy.

Table 5.2 Player 1's payoffs given mixed strategy (two-player experiment, first price auction)

Bid price		Player 2			
		10	15	20	25
Player 1	10	$\frac{10 - Cost_1^t}{2}$	$10 - Cost_1^t$	$10 - Cost_1^t$	$10 - Cost_1^t$
	15	0	$\frac{15 - Cost_1^t}{2}$	$15 - Cost_1^t$	$15 - Cost_1^t$
	20	0	0	$\frac{20 - Cost_1^t}{2}$	$20 - Cost_1^t$
	25	0	0	0	$\frac{25 - Cost_1^t}{2}$

Table 5.3 Player 2's payoffs given mixed strategy (two-player experiment, first price auction)

Bid price		Player 1			
		10	15	20	25
Player 2	10	$\frac{10 - Cost_2^t}{2}$	$10 - Cost_2^t$	$10 - Cost_2^t$	$10 - Cost_2^t$
	15	0	$\frac{15 - Cost_2^t}{2}$	$15 - Cost_2^t$	$15 - Cost_2^t$
	20	0	0	$\frac{20 - Cost_2^t}{2}$	$20 - Cost_2^t$
	25	0	0	0	$\frac{25 - Cost_2^t}{2}$

The player with the lowest bid price wins the game in the first price auction. For the three-player experiment and second price auction, the payoff is determined by the second lowest bid and the player's given cost in that game. Therefore, players in the second price auction can expect a greater payoff than the players in the first price auction. The payoff function for the second price auction game is as follows:

$$\pi_n^{it} = \begin{cases} SBP_m^{jt} - Cost_n^t & \text{If each player chooses the different bid price alternative} \\ \left(\dfrac{SBP_m^{jt} - Cost_n^t}{2} \right) & \text{Otherwise} \end{cases}$$

where

π_n^{it} : player n's payoff for choosing the alternative bid price i at time (game) t

SBP_n^{it} : player m's bid, which is the second lowest price, as alternative j at time (game) t

$Cost_n^t$: player n's cost at time (game) t

5.3. Two-Player Experimental Data Analysis

Figures 5.3 and 5.4 show the player's actual frequency of bid price choices for two-player experiment, given epistemic learning data. The actual frequency ($F_n^t(i)$) of bid price alternative i is calculated by:

$$F_n^t(i) = \frac{N_n^t(i)}{\sum_j^{C_n} \sum_t N_n^t(j)} \tag{5.2}$$

$F_n^t(i)$: the player n's actual frequency for choosing alternative bid price i at time (game) t

$N_n^t(i)$: the player n's total number of choice for choosing alternative bid price i at time (game) t

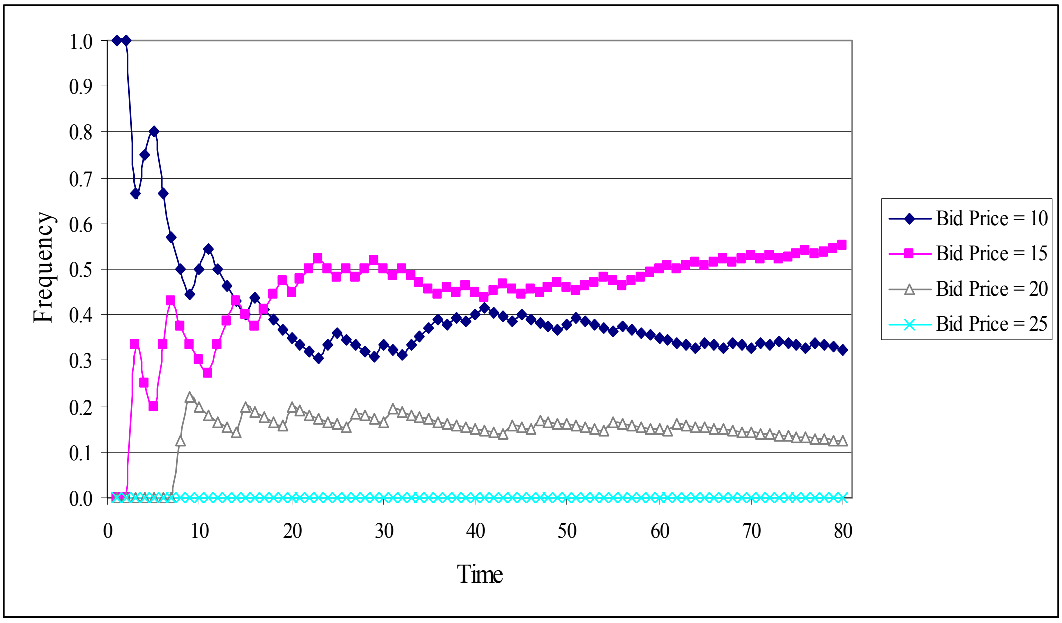


Figure 5.3 Player 1's actual frequency of bid price choice for epistemic learning data (two-player experiment)

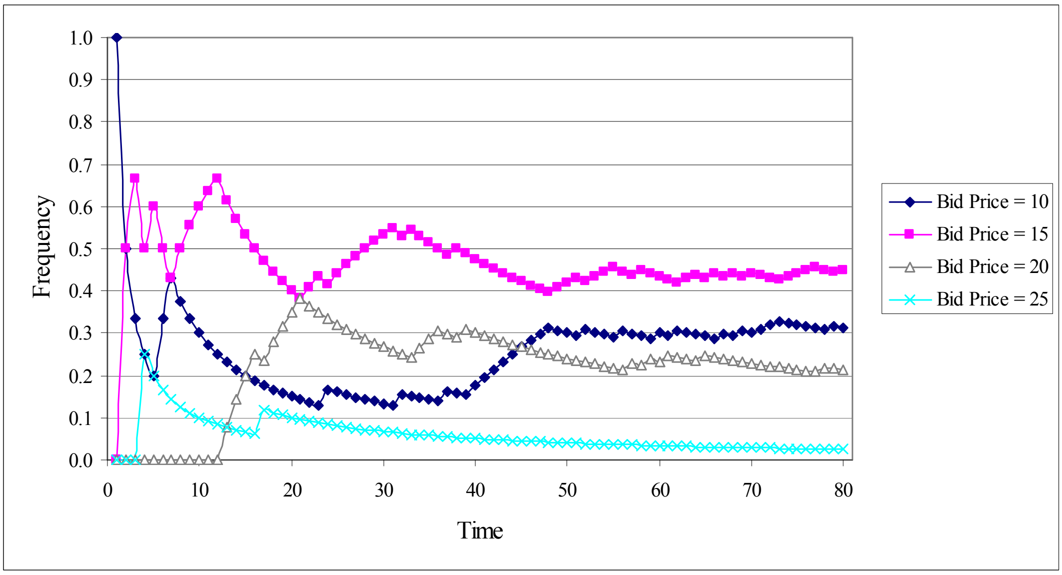


Figure 5.4 Player 2's actual frequency of bid price choice for epistemic learning data (two-player experiment)

Figures 5.3 and 5.4 illustrate that the actual frequency of bid price 15 for both players, using epistemic learning data, is higher than for all alternative choices after ten to twenty games (*t*=10~20). As games are repeated, Player 1's actual frequency of choosing bid price 15 (*BP*=15) increases relative to the other bid price choices. However, as games are repeated, Player 2's frequency of bid price 10 (*BP*=10) dramatically increases after 40 games (*t*=40), but the frequency of bid price 15 (*BP*=15) remains largely unchanged after 45 games. As mentioned before, the players have knowledge about their opponents' previous choices (actions). Player 2 prefers to choose alternative 10 (*BP*=10) after 40 games, because he or she learned that Player 1 is more likely to choose bid price 15 (*BP*=15) than any other choice. These patterns are depicted clearly in Figures 5.4 and 5.9. In Figure 5.9, Player 2's probability of winning is increased after 40 games (*t*=40), since Player 2 chooses the lowest bid price 10 more often than any other choices.

Figure 5.5 presents the players' probability[13] of each alternative over 80 games, using the epistemic learning experimental data. The probability of each bid price choice for each player shows the player's bidding strategy and his or her own risk management preference. From this perspective of risk management, Player 2 is a greater risk taker than Player 1. Player 2 is more likely to choose bid price 20 or 25 than Player 1, if he or she can have more payoffs.

---

[13] $F_n^{80}(i) = \dfrac{N_n^{80}(i)}{\sum_{j=1}^{4}\sum_{t=1}^{80} N_n^t(j)}$

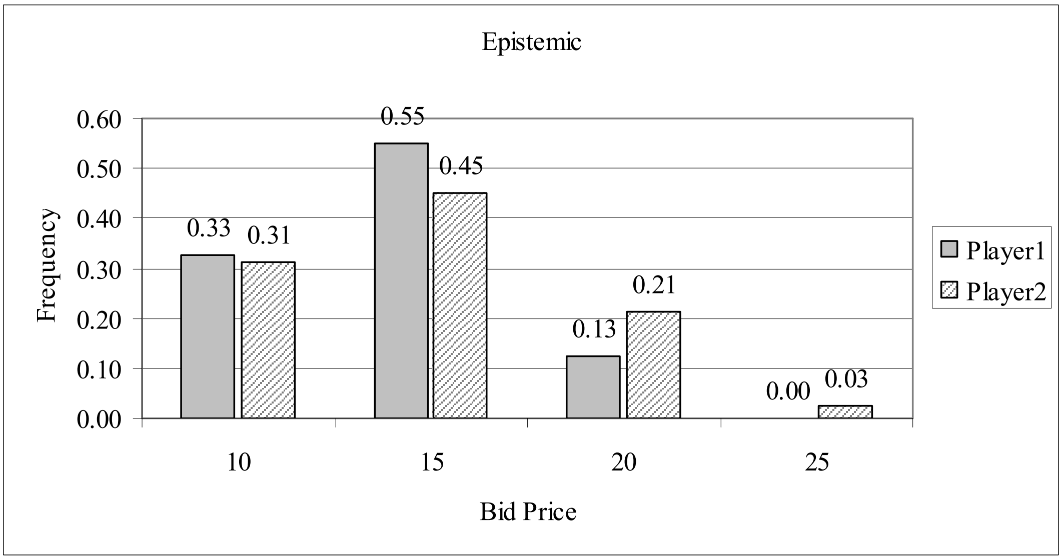


Figure 5.5 Probability of choice over 80 games for epistemic learning (two-player experiment)

We estimate a linear regression model, with bid price as the dependent variable, to test whether a winning payoff variable has a significant effect on bid price choice. The following linear regression structure is examined to show the relationship between the chosen bid price and payoffs, if a player wins that game. Table 5.4 presents the regression results for the epistemic learning experimental data.

$$BP_n^t = \alpha + \beta \times payoff_n^t \quad (5.3)$$

where

α : a constant

β : the coefficient of winning payoff variable

BP_n^t : player n 's chosen bid price at time (game) t

$payoff_n^t$: player n 's payoff at time (game) t if player wins that game

Table 5.4 Regression results for epistemic learning data (two-player experiment)

	Player1		Player2		Both	
	B	t-value	B	t-value	B	t-value
(Constant)	14.976	40.545	15.061	41.725	15.008	60.535
Payoff	0.789	5.049	0.753	6.597	0.767	8.637
R^2	0.246		0.358		0.321	
SSE	618.032		799.146		1418.028	
SSR	201.968		445.854		669.472	
SST	820.000		1245.000		2087.500	

As can be seen from Table 5.4, all variables have significant t-values at a 95% confidence level. This means that the winning payoff effect on bid price choice is significant. It also shows that Player 2 is more concerned about the winning payoff to choose bid price than Player 1, as seen by the R-square value. However, the R-square values for *goodness of fit* are not high enough to explain either player's bidding behavior. In particular, the interaction between players affects each player's decisions in the epistemic learning experiment. This implies that a player cannot choose a bid price alternative based upon the winning payoff.

Figures 5.6 and 5.7 present the actual frequency of bid price choices for the behavioral reinforcement experiment data. Figure 5.8 shows the probability of each bid price choice across 80 games. In the behavioral reinforcement experiment, players only know their own actions. The rank order of Player 1's frequency of bid price choices at the end of sequential games is the same as the corresponding rank order observed in the epistemic learning experiment. On the other hand, Player 2 is more likely to choose bid price 10 (BP=10) than bid price 15 (BP=15). After 40 games, Player 1 prefers to choose bid price 15 (BP=15), but Player 2 is more likely to choose bid price 10 (BP=10). Therefore, Player 2's winning probability is higher than Player 1's winning probability after 40 games (see Figure 5.10).

As mentioned before, the probability of each bid price choice for each player, shown in Figure 5.8, demonstrates each player's bidding strategy and his or her own risk management preference. Player 1 takes greater risks in the behavioral reinforcement game than in the epistemic learning game. However, Player 2 behaves in a more risk neutral manner, by choosing bid prices 10 and 15 at almost the same frequency.

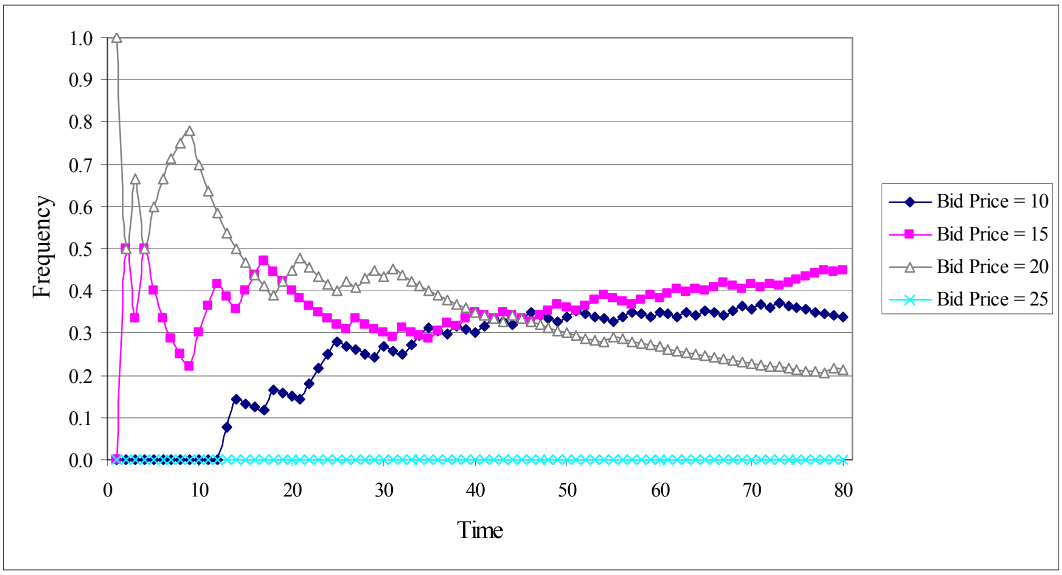


Figure 5.6 Player 1's actual frequency of bid price choice for behavioral reinforcement learning data (two-player experiment)

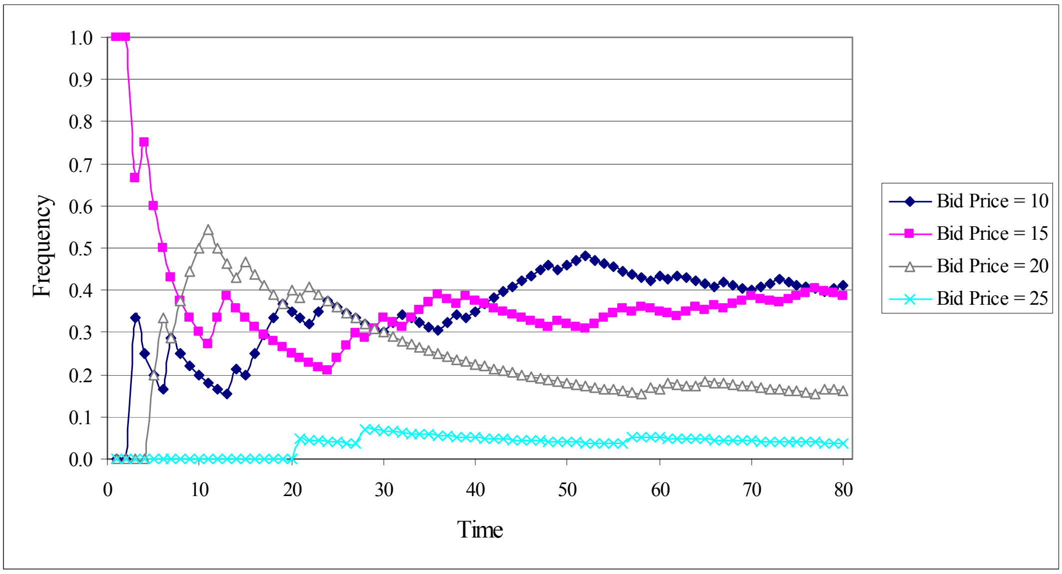


Figure 5.7 Player 2's actual frequency of bid price choice for behavioral reinforcement learning data (two-player experiment)

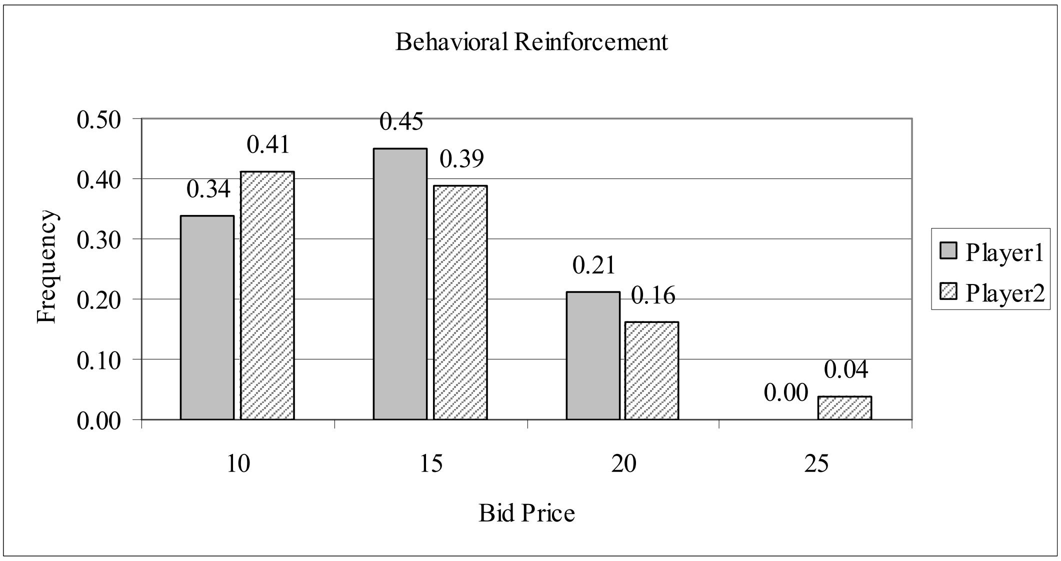


Figure 5.8 Probability of choice over 80 times for behavioral reinforcement (two-player experiment)

A linear regression model, with bid price as the dependent variable, was constructed to test whether a winning payoff variable has a significant effect on bid price choice. The following linear regression structure, using behavioral reinforcement experimental data, shows the relationship between chosen bid price and payoff, if a player wins that game (see Table 5.5).

Table 5.5 Regression results for behavioral reinforcement learning data (two-player experiment)

	Player1		Player2		Both	
	B	t-value	B	t-value	B	t-value
(Constant)	15.076	46.359	15.088	38.499	15.075	59.653
Payoff	0.813	7.688	0.928	7.099	0.868	10.434
R^2	0.431		0.393		0.408	
SSE	608.029		843.615		1456.443	
SSR	460.721		545.135		1003.557	
SST	1068.750		1388.750		2460.000	

The winning payoff variable has a significant t-value at a 95% confidence level for all cases. The R-square value for the behavioral reinforcement game is higher than for the epistemic learning game, since players only consider their own payoffs in this game. Distinct from the epistemic result, Player 1 is more concerned about the winning payoff when choosing bid price than Player 2, as seen by the R-square value.

Figures 5.9 and 5.10 present players' winning and tying probabilities for each experiment. Table 5.6 shows the number of ties and probability of bid price choices for each experiment.

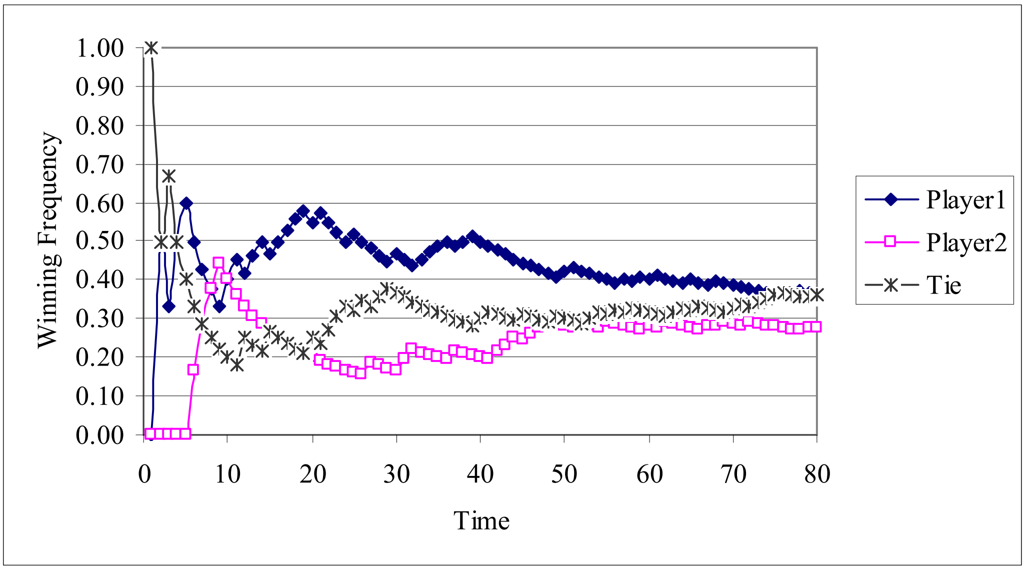


Figure 5.9 Players' winning frequency[14] for epistemic learning data (two-player experiment)

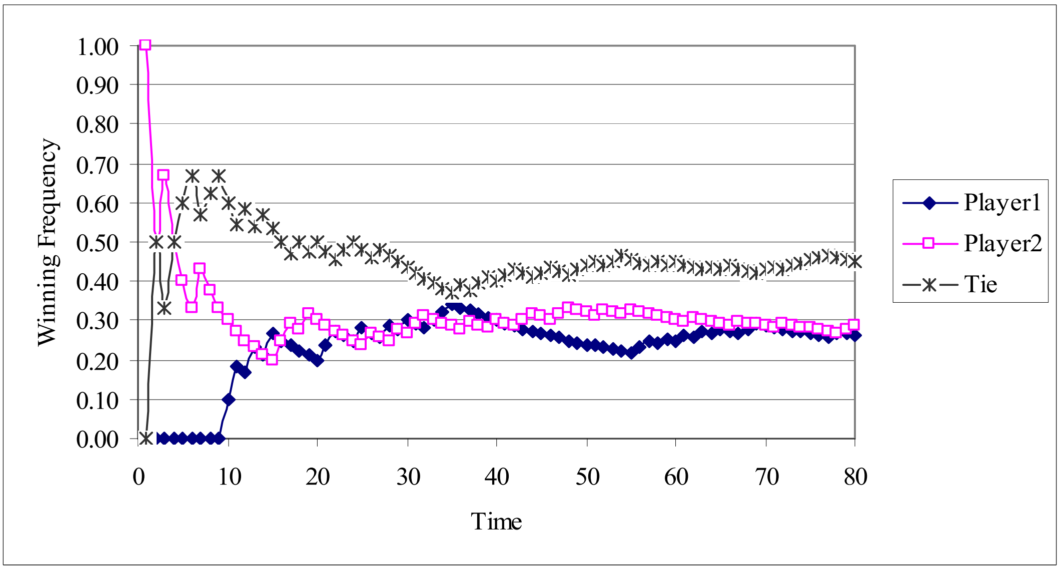


Figure 5.10 Players' winning frequency for behavioral reinforcement learning data (two-player experiment)

---

[14] Winning frequency includes the number of tie game

Table 5.6 Frequency of tie for epistemic and behavioral reinforcement learning (two-player experiment)

	Epistemic		Behavioral Reinforcement	
Bid Price	# of Tie	%	# of Tie	%
10	8	27.6	14	38.9
15	19	65.5	17	47.2
20	2	6.9	5	13.9
25	0	0.0	0	0.0
total	29	100.0	36	100.0

As a player chooses a lower bid price, a player achieves less as a payoff. In the epistemic learning experiment, each player knew who had chosen which bid price after each game. Accordingly, each player was very sensitive to the other's actions. The probability of a tie is less in the epistemic learning game than in the behavioral reinforcement learning game (see Figures 5.9 and 5.10, and Table 5.6). This implies that, in the epistemic learning game, players choose different bid prices, because players can predict other player's actions. For this reason, the number of ties with bid price 15 (BP=15) is much higher than the other choices in the epistemic learning game (see Table 5.6). In the behavioral reinforcement experiment, players cannot predict the other player's actions; they only know the opponent's choice if they bid the same price. Therefore, ties occur with bid prices 10 and 15 more often than with any other bid price.

5.4. Three-Player Experimental Data Analysis

The second experiment was conducted to evaluate the performance of the dynamic strategy choice model. In these experiments, three players attended 80 games, most of which were the same as in the first experiment (except for the price award assumption). Players were notified that the game winner would receive a payoff based not on their bidding price but on the second lowest bid. We expected that players in this second price auction would bid lower than the players in the first price auction. Accordingly, the player's reaction against other competitors' strategies should differ compared to the players' behavior in the first price auction. In addition, it was difficult for players to determine their bid, since players did not know their true value of their choice. In classic second price auctions, players often fail to adjust their value estimates to the information revealed by winning; hence, players have less confidence in their understanding of the auction and prefer to avoid behavior that appears extreme (Kagel, 1995). Under such conditions, the results of this second experiment should differ from those of the first experiment. The following analysis for the three-player experiment demonstrates each player's preference for each strategy and their bidding behavior for each cognitive learning process.

Figures 5.11 through 5.13 show the player's actual frequency of bid price choices, given epistemic learning data.

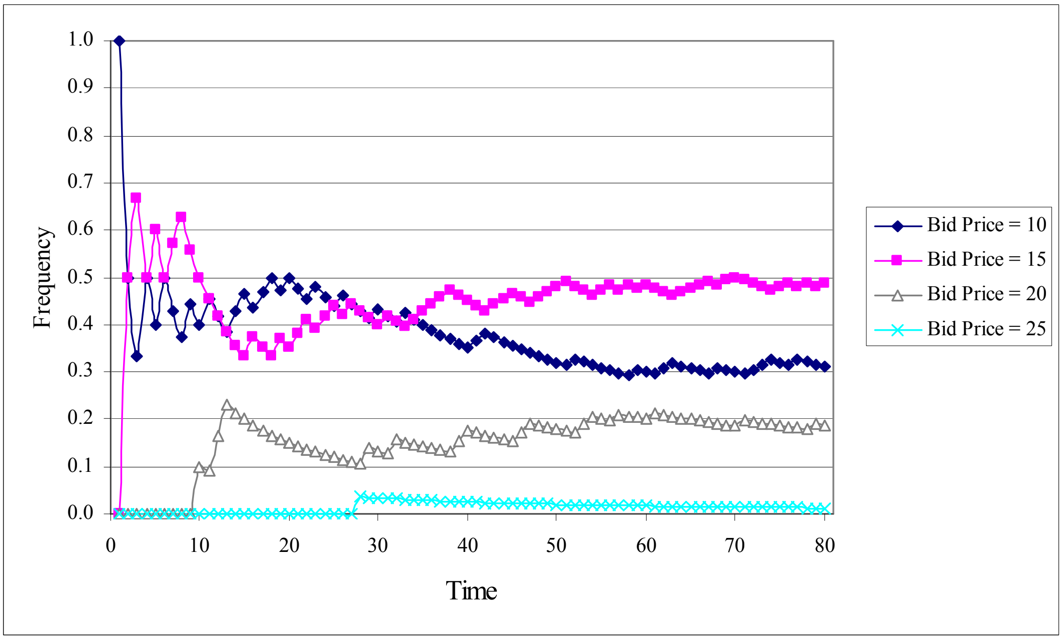


Figure 5.11 Player 1's actual frequency of bid price choice for epistemic learning data (three-player experiment)

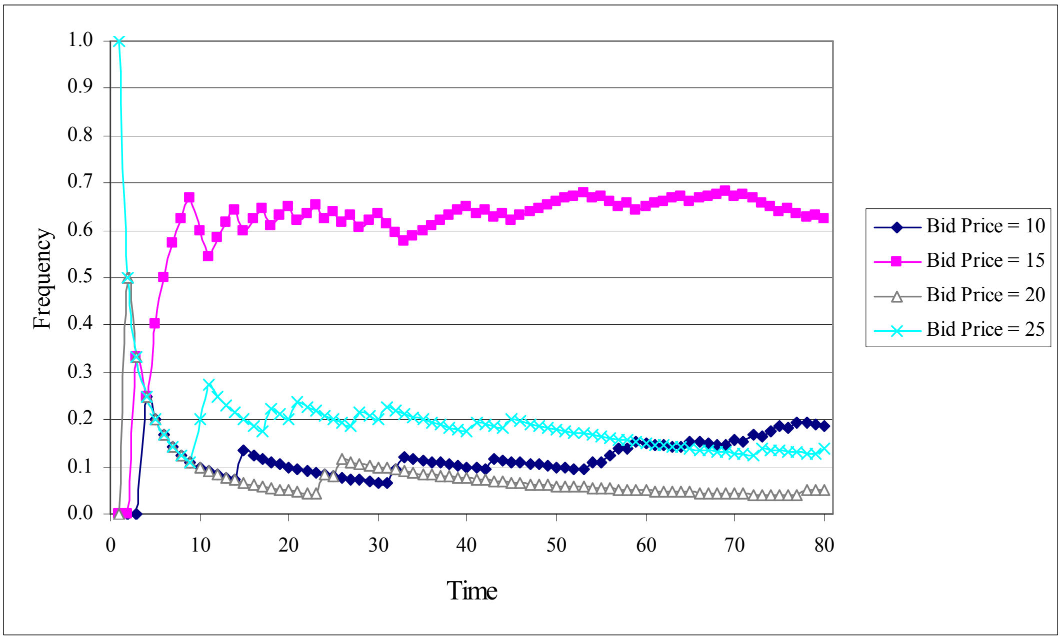


Figure 5.12 Player 2's actual frequency of bid price choice for epistemic learning data (three-player experiment)

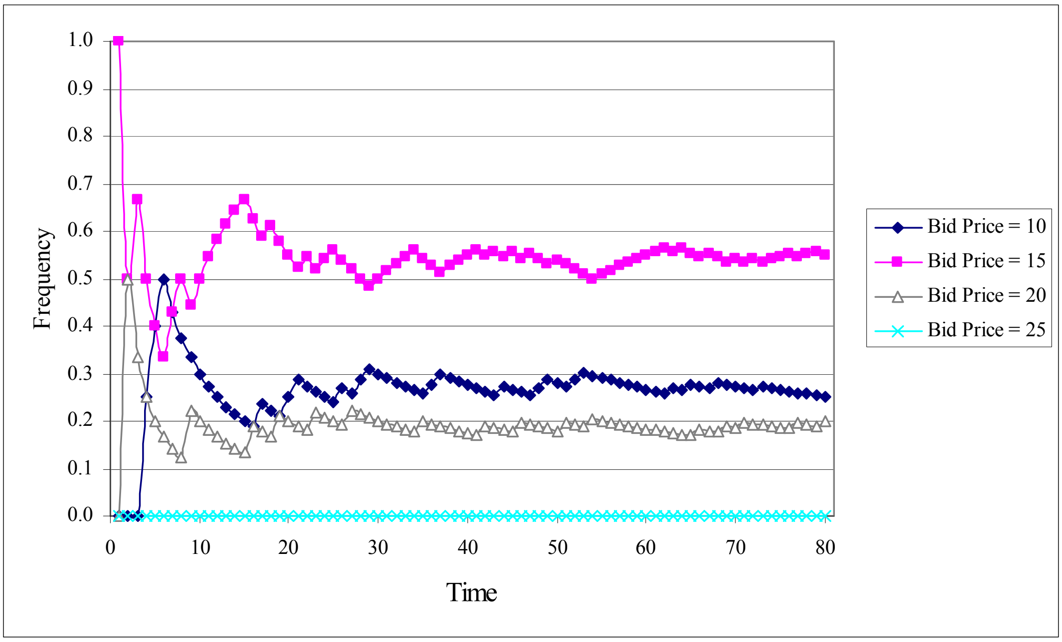


Figure 5.13 Player 3's actual frequency of bid price choice for epistemic learning data (three-player experiment)

Figures 5.11 through 5.13 show that the actual frequency of bid price 15 (BP=15) for all players is higher than for other choices after 10 to 25 games (t=10–25). The rank order of bidding frequency at the end of sequential games is the same as the corresponding order observed in the two-player experiment. As games repeated, the actual frequency of choosing bid price 15 (BP=15) among all players increased relative to the other bid price choices. In particular, this bid was a dominant choice for Players 2 and 3 in all games. Players' payoffs for bidding a price of 10 was mostly negative, since the given cost range was between 10 and 20. Players could not take the risk of bidding 10, even in the second price auction game, since they had full information of other player's bidding history. However, in the second price auction game, the winning player's payoff is the difference between the second lowest price and given player's own cost. Therefore, the frequency of bid price 15 increased overall for all games compared to prices 20 and 25, since players can expect more payoffs and opportunity to win that game by choosing a bid price of 15.

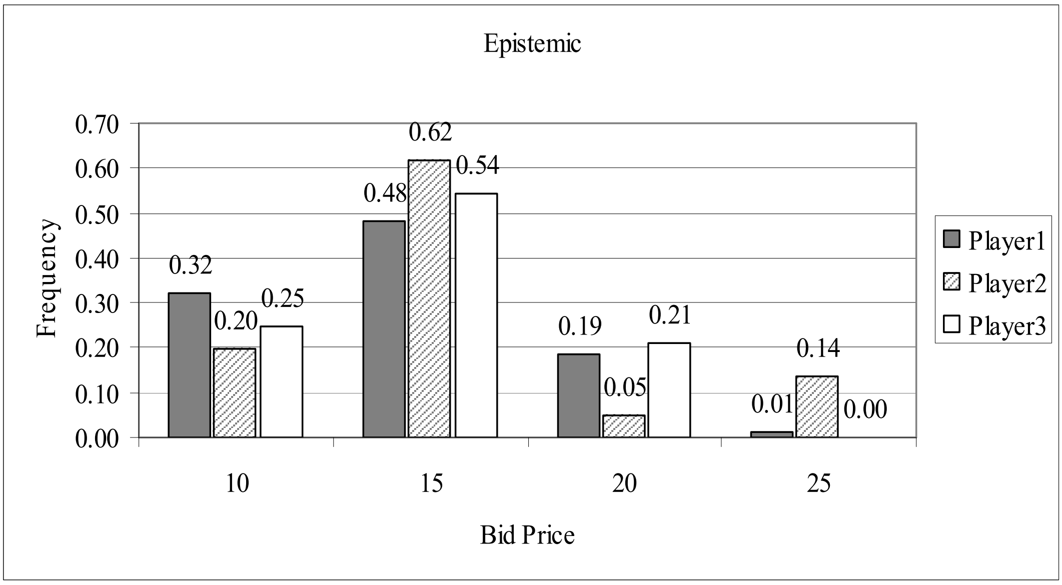


Figure 5.14 Probability of choice over 80 games for epistemic learning game (three-player experiment)

Figure 5.14 indicates the players' probability[15] of selecting each alternative over the 80 games, using the epistemic learning data, as well as each player's risk management preference. From this perspective, all players were risk neutral compared to those in the first experiment. Players were more competitive in this second experiment. Therefore, because all players had information about competitors'

---

[15] $F_n^{80}(i) = \dfrac{N_n^{80}(i)}{\sum_{j=1}^{4}\sum_{t=1}^{80} N_n^t(j)}$

history, they could not afford to take the same risks as those players in the first experiment.

A linear regression model is estimated using the epistemic learning data. The bid price is a dependent variable, and the winning payoff is an independent variable. In the second auction game, this winning payoff used in a regression model is not an exact payoff that players can have after that game if they win that game. However, we use this payoff as an independent variable, since players cannot predict the real values of their payoffs. They expect that they will have more payoffs than this winning payoff, which is the difference between the player's bid price and given costs, if they win that game. This model tests whether a winning payoff variable has a significant effect on the choice of bid price. The linear regression formulation (equation (5.3)) shows the relationship between the chosen bid price and the payoffs received. Table 5.7 presents the regression results for the epistemic learning experimental data.

Table 5.7 Regression results for epistemic learning data (three-player experiment)

	Player1		Player2		Player3		All	
	B	t-value	B	t-value	B	t-value	B	t-value
(Constant)	14.802	43.550	15.414	48.685	15.328	44.212	15.203	80.257
Payoff	0.924	6.462	0.971	10.721	0.858	4.987	0.943	13.469
R^2	0.346		0.593		0.239		0.429	
SSE	719.623		639.181		701.389		2082.166	
SSR	380.377		929.954		220.833		1567.423	
SST	1100.000		1569.136		922.222		3649.588	

We found a significant positive correlation between payoffs and bid price based on the t-values at a 95 percent confidence level. Players having more winning payoffs bid higher prices. The Table 5.7 shows that Player 2 is more concerned about the winning payoff than other players in the epistemic learning game, as indicated by the R-square value. This R-square value indicates that 59.3 percent of the variance in bid price can be predicted from the player's winning payoffs variable. Note that this is an overall measure of the strength of association; however, the R-square values for *goodness-of-fit* are not high enough to explain either player's bidding behavior. In the multi-player and the second price auction games, players could consider the other competitor's strategies by observing their sequence of choice. Thus, more related factors must be included in the model structure.

Figures 5.15, 5.16, and 5.17 present the actual frequency of bid price choices for each player using behavioral reinforcement data. In this experiment, players could track their own history of bidding choices over several games. The rank order of player's frequency of bid price choices at the end of sequential games differs from the corresponding rank order observed in the three-player epistemic learning experiment. Players are more likely to choose bid price 15 (BP=15) than other prices. Accordingly, players prefer to choose bid price 20 (BP=20) over bid prices 10 (BP=10) and 25 (BP=25). Player 3 prefers to choose bid price 20 over bid price 15 until 35 games (t=0~35). This example shows that players try to avoid choosing between two extreme bid prices (highest and lowest), since players may have low payoffs and a low probability to win. In this experiment, the second price auction was applied, which can be riskier for players, since players can rely on the auction type and their payoffs in the behavioral reinforcement learning game. In the epistemic learning game, players cannot take as many risks as those in the behavioral reinforcement game, since players know each other's choices. Therefore, the probability of choosing bid price 10 in the epistemic learning game for the three-player experiment is higher than choosing bid prices 20 or 25 (see Figures 5.11 through 5.13).

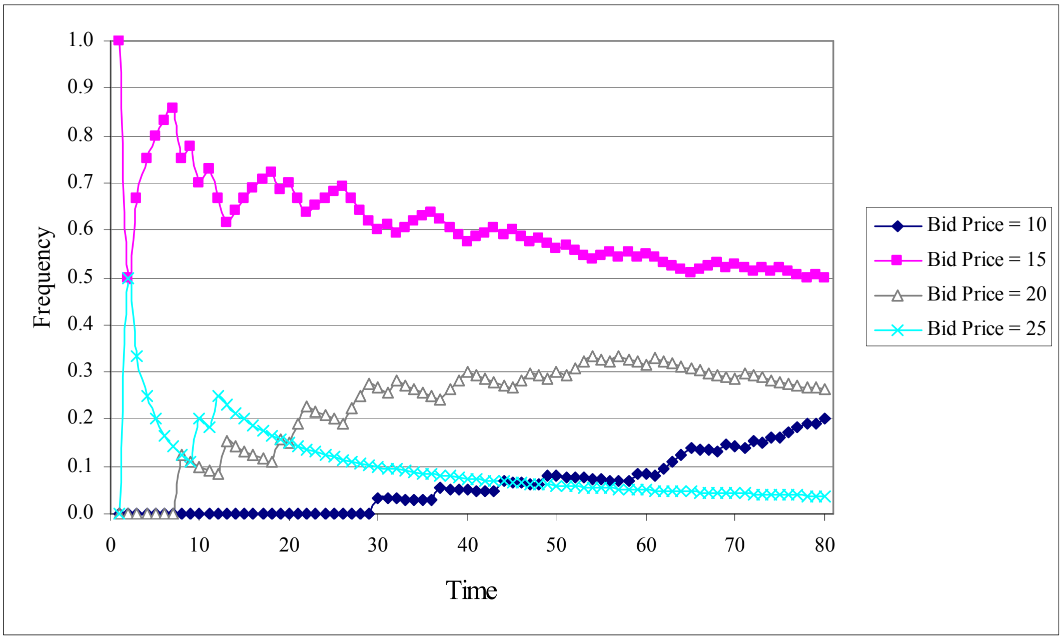


Figure 5.15 Player 1's actual frequency of bid price choice for behavioral reinforcement learning experiment data (three-player experiment)

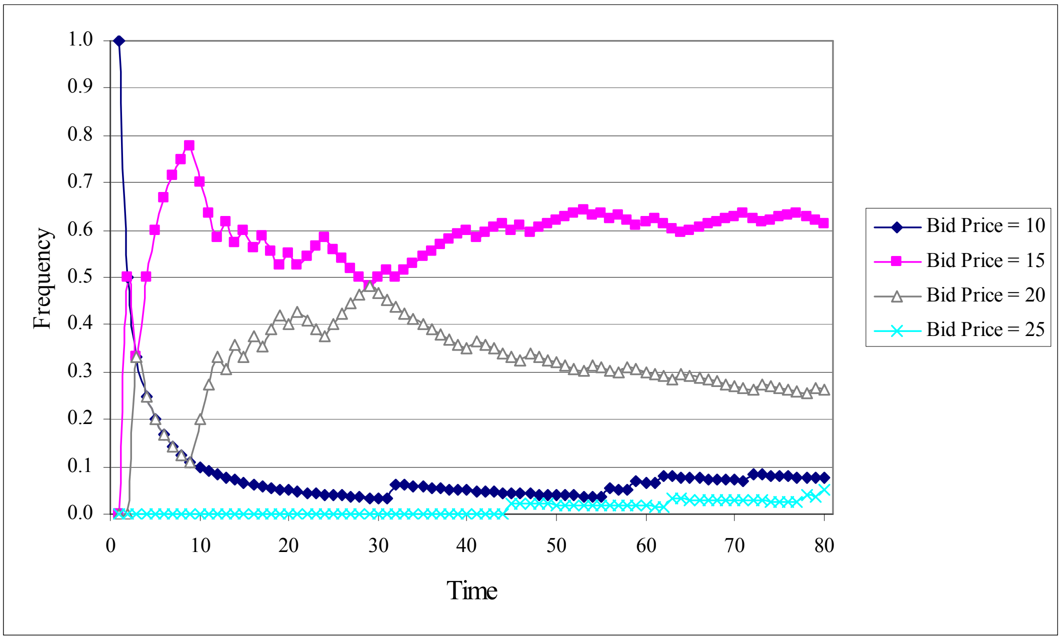


Figure 5.16 Player 2's actual frequency of bid price choice for behavioral reinforcement learning experiment data (three-player experiment)

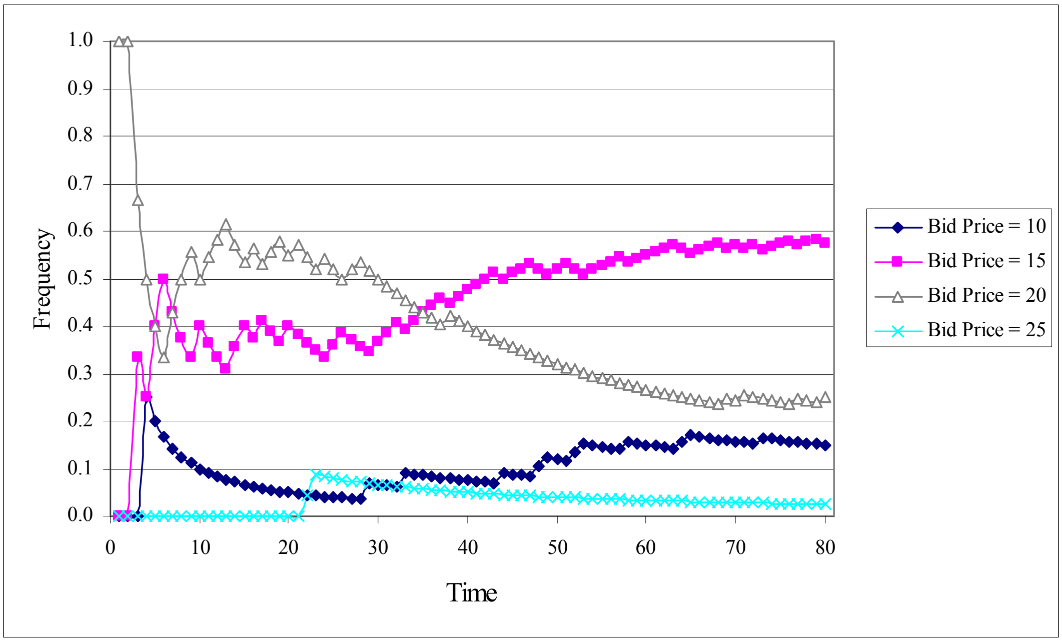


Figure 5.17 Player 3's actual frequency of bid price choice for behavioral reinforcement learning experiment data (three-player experiment)

Figure 5.18 shows the probability of each bid price choice across 80 games using behavioral reinforcement data. The probability of each bid choice reveals each player's bidding strategy and risk management preference. Player 2 chooses bid price 10 less often than other players. Player 2 takes greater risks in the behavioral reinforcement experiment than other players. By choosing bid prices 10 and 15, Players 1 and 3 are more risk neutral compared to Player 2.

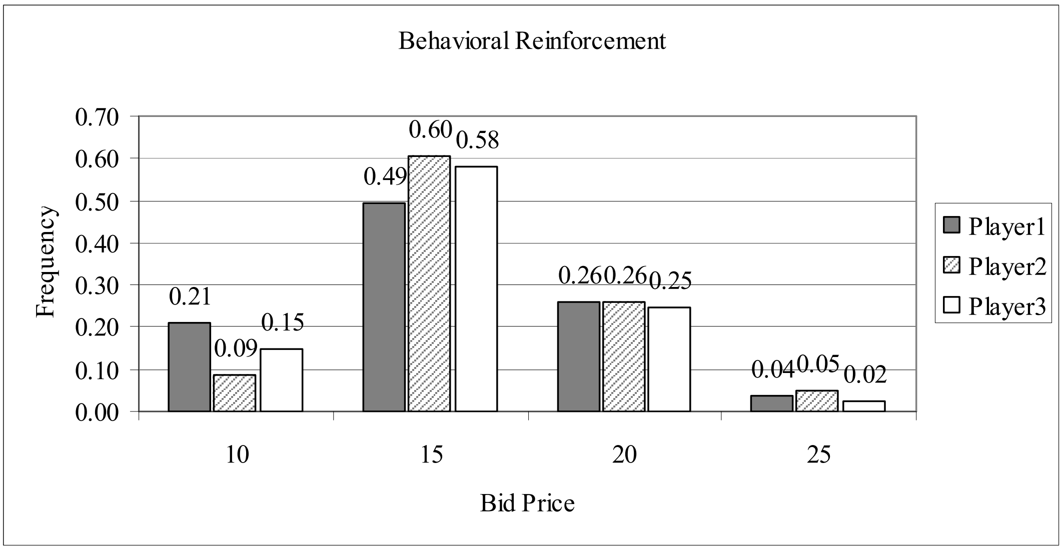


Figure 5.18 Probability of choice over 80 times for behavioral reinforcement (two-player experiment)

Table 5.8 shows the linear regression model estimation results using behavioral reinforcement learning data. We try to determine the relationship between two variables: bid price and winning payoffs.

Table 5.8 Regression results for behavioral reinforcement learning experiment data (three-player experiment)

	Player1		Player2		Player3		All	
	B	t-value	B	t-value	B	t-value	B	t-value
(Constant)	14.886	42.583	15.657	48.691	15.527	49.999	15.384	81.658
Payoff	0.931	7.407	0.738	6.870	0.654	6.682	0.759	12.033
R^2	0.410		0.374		0.361		0.375	
SSE	719.493		595.107		610.510		1968.314	
SSR	499.643		355.510		345.045		1182.509	
SST	1219.136		950.617		955.556		3150.823	

Table 5.8 presents the value of the R-square, which is less than 0.5. This means that this model is not a good reflection of the relationship between bid price and winning payoffs. The winning payoff variable has a significant t-value at a 95 percent confidence level for all cases. This demonstrates that our estimates are still statistically unbiased but are infected with more noise. Unlike with the epistemic results, Player 1 is more concerned about the winning payoff when choosing a bid price than Players 2 and 3, as seen by the R-square value. From the above results, we can conclude that the winning payoff cannot explain each player's bidding behavior; such relationships cannot be linearly described. Therefore, we consider different types of variables, the average payoff for behavioral reinforcement learning, and possible payoffs based on the player's mixed strategy choice for epistemic learning to describe bidding behavior.

Figures 5.19 and 5.20 show the probability of winning or tying for each experiment. In the epistemic learning game, players received information about other players' choices after each game. Therefore, each player could make a decision after considering other players' actions. Accordingly, the frequency of ties for players in the epistemic learning game was less than in the behavioral reinforcement game. The epistemic learning game in Figure 5.19 shows that the winning frequency for Player 1 is higher than for other players in all 80 games; Player 1 is a dominant winner. At the end of game, the frequency of winning for all players is very similar. For the two-player experiment, on the other hand, there was a dominant winner in the epistemic learning game (see Figure 5.9). This implies that a player with more winning experience can retain winning probability, since that winner has a greater ability to predict other player's bidding behavior using epistemic learning information.

In the behavioral reinforcement game, however, we could not say which player was the dominant winner. The behavioral reinforcement learning process assumes that players cannot track competitor's choices; they can only review their own past payoffs before making a decision. Therefore, no player can predict their competitor's next action. A player whose bid choice resulted in high payoffs in the past will have more payoffs.

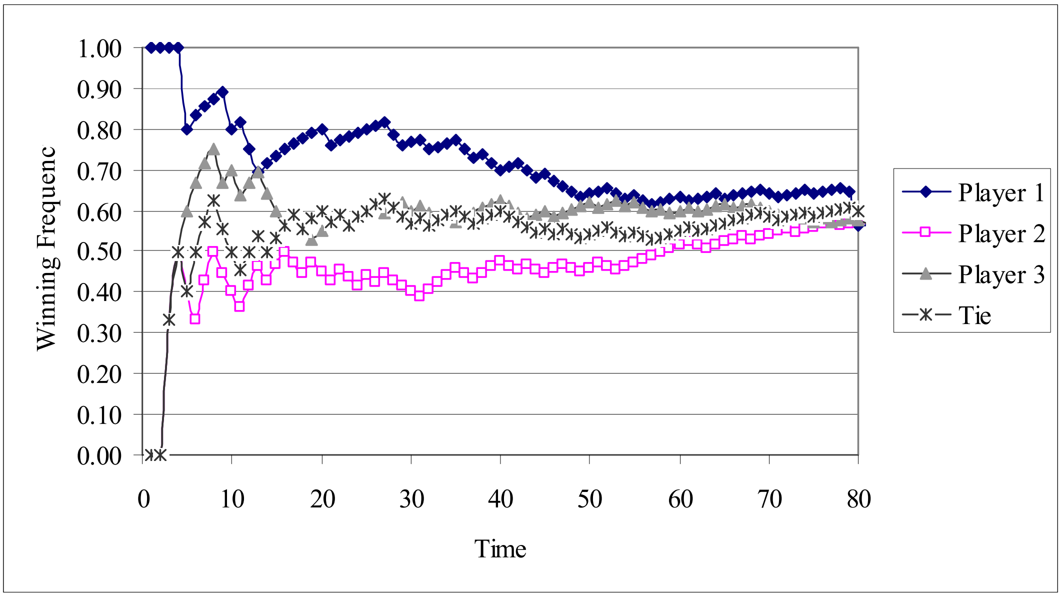


Figure 5.19 Players' winning frequency[16] for epistemic learning experiment data (three-player experiment)

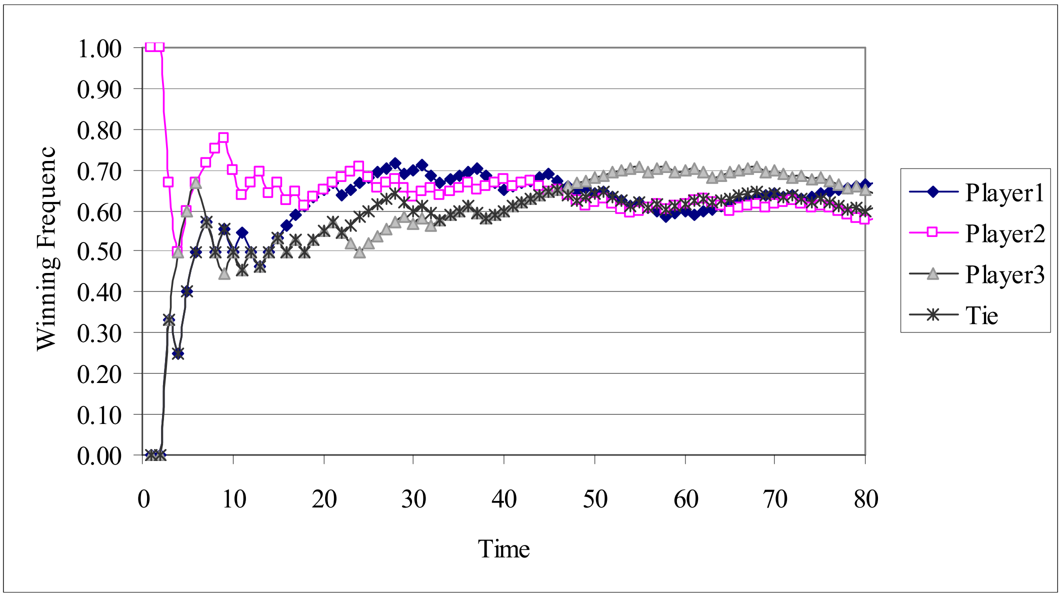


Figure 5.20 Players' winning frequency for behavioral reinforcement learning experiment data (three-player experiment)

---

[16] Winning frequency includes the number of tie game

Chapter 6. Dynamic Multinomial Probit Estimation Results

6.1. Introduction

The previous chapter discussed the results of the explanatory analysis on two experimental datasets. This chapter focuses on providing empirical results and estimates of the two models, using data from the two sequential choice experiments described in Chapter 5. Before the estimation, we must classify the two learning models corresponding to the two experimental datasets and the two types of mixed learning model corresponding to the epistemic experimental datasets. In addition, three types of learning model specifications (epistemic, behavioral reinforcement, and mixed learning) are investigated for DMNP and DML model frameworks. Three different models of learning procedures are estimated to further demonstrate the capability of the developed MLE estimation procedure for DMNP models with a large number of parameters, alternatives, and error structures.

We interpret the estimation results for cognitive learning models by evaluating the performance of the learning sequences of choices mentioned earlier in the DMNP model. In this chapter, we first discuss the epistemic learning estimation results generated by the DMNP model (Section 6.2); then, we present the behavioral reinforcement learning estimation results (Section 6.3); In Section 6.4, the statistical test results are presented to test the mis-specification for both learning models; finally, we specify the mixed learning models and provide the estimation results of the mixed learning model (Section 6.5). All numerical experiments in this study were implemented using FORTRAN as the programming language.

To simplify the estimate, the same variance and covariance were assumed in v_n^{it} and $\Omega_n^{it} : \sigma_i^2 = \sigma_j^2$, $\gamma_{nm}^{iit} = \gamma_{nm}^{jjt}$, $\gamma_{nm}^{ijt} = \gamma_{nm}^{jit}$, and $\gamma_{nm}^{ijt} = \gamma_{nm}^{ijt'}$. The estimation results for the behavioral reinforcement, epistemic, and mixed learning models are shown, with error structure in the following sections.

6.2. Epistemic Learning Behavioral Interpretation

The explanatory variables included in the empirical epistemic learning model specification (Table 6.1) are the player's expected payoffs given their beliefs[17] regarding the opponent's type and habit persistence. Average payoff, cumulative choice sequence, and habit persistence variables are used in the behavioral reinforcement learning model specification in Tables 6.2 and 6.3.

Table 6.1 presents the estimation results using the datasets for the epistemic learning case. As expected, the coefficients for the player's expected payoffs, given a belief regarding the opponent's type in the epistemic learning model, is positive and strongly significant. The empirical results for the epistemic learning model indicate that, given a belief regarding the opponent's type, players who feel more positively about the expected payoff have a greater utility for the alternative bid price than

[17] the form of stationarity of opponents' behavior

players who have less expected payoff. These results imply that a player who follows the epistemic learning rule is sensitive to other players' choice decisions, since mixed strategies given by players' choice combinations determine the payoffs. However, the results also demonstrate the decreasing utility of bid price with increasing utility of previous time periods in habit persistence term. A player adopting the epistemic learning rule is more concerned with other players' current possible choices than with his or her own previous utility. Players tend to believe that the previous high utility of that bid price choice cannot guarantee the high utility of that bid price in the current game. Players prefer to know the opponent's type before making a decision each game.

In order to test the player's epistemic learning behavior, epistemic learning datasets were applied to the behavioral reinforcement model specification and the estimation results are presented in Table 6.2. The hypothesis of this experiment is that even though a player has knowledge of an opponent's type, the player will behave in a manner consistent with the behavioral reinforcement learning rule. The results in Table 6.2 show that the coefficients for the explanatory variables and error terms are significant, but the log likelihood ratio in Table 6.2 is less than that of Table 6.1 (the epistemic learning model). This may imply that players who hold to the epistemic learning rule are more likely to use all information about their opponent's type to make a decision.

Table 6.1 Estimated parameters of the epistemic learning DMNP model using epistemic learning data for the two-player experiment

	Epistemic Learning Model[18]	
Attributes	Estimates	T-value
Initial Value (β_I)	-0.842	-5.836
Constant(alt1, β_0^1)	-0.087	-8.720
Constant(alt2, β_0^2)	-0.712	-3.369
Constant(alt3, β_0^3)	-0.507	-4.412
Payoff based on the belief of opponents action (β_1)	1.770	4.505
Habit persistence 1 (β_2)	-0.288	-3.296
ρ_1	0.408	4.823
ρ_2 (Habit persistence 2)	0.369	9.950
γ_{12}^{12}	0.356	2.634
γ_{12}^{13}	0.481	2.006
γ_{12}^{14}	-0.005	-3.568
γ_{12}^{23}	0.003	7.173
γ_{12}^{24}	-0.003	-1.639
γ_{12}^{34}	0.029	4.500
γ_{21}^{12}	-0.594	-1.989
γ_{21}^{13}	-0.008	-7.579
γ_{21}^{14}	0.853	1.302
γ_{21}^{23}	-1.241	-4.569
γ_{21}^{24}	1.162	4.144
γ_{21}^{34}	-0.081	-2.810
Log-likelihood at convergence	-0.4175	

[18] Epistemic Learning Model:

$$U_n^{it} = \beta_I Y_n^{it} + \beta_0^i (1 - Y_n^{it})$$

$$+ \beta_1 \sum_{k=1}^{C_{-n}} \pi_n(s_n^i, s_{-n}^k(t-1)) \cdot p(s_{-n}^k(t) \middle| s_{-n}^k(t-q), q = 1,\ldots,t-1) + \beta_2 V_n^i(t-1)(1 - Y_n^{it}) + \varepsilon_n^{it}$$

$$p(s_{-n}^k(t) \middle| s_{-n}^k(t-q), q = 1,\ldots,t-1) = \frac{N_{-n}^k(t-2) + I(s_{-n}^k(t-1))}{\sum_{h=1}^{C_{-n}} \left[N_{-n}^h(t-2) + I(s_{-n}^h(t-1)) \right]}$$

Table 6.2 Estimated parameters of the behavioral reinforcement learning DMNP model using epistemic learning data for the two-player experiment

	Epistemic learning data to Behavioral Reinforcement Model[19]	
Attributes	Estimates	T-value
Initial Value (β_I)	0.000	-7.777
Constant(alt1, β_0^1)	-0.207	-4.996
Constant(alt2, β_0^2)	0.671	5.842
Constant(alt3, β_0^3)	0.408	3.343
Avg. payoff (β_1)	0.490	3.539
Sequence of choice (β_2)	0.517	3.374
Habit persistence 1 (β_3)	-0.140	-2.056
ρ_1	0.469	11.335
ρ_2 (Habit persistence 2)	-0.136	-4.258
γ_{12}^{12}	0.131	2.693
γ_{12}^{13}	0.145	3.147
γ_{12}^{14}	0.000	-5.519
γ_{12}^{23}	-0.429	-1.958
γ_{12}^{24}	-0.076	-3.930
γ_{12}^{34}	0.500	8.144
γ_{21}^{12}	-0.484	-2.609
γ_{21}^{13}	0.207	3.247
γ_{21}^{14}	0.065	9.365
γ_{21}^{23}	-0.334	-1.844
γ_{21}^{24}	0.074	13.652
γ_{21}^{34}	0.290	2.791
Log-likelihood at convergence	-0.9039	

---

[19] Behavioral Reinforcement Learning Model:

$$U_n^{it} = \beta_I Y_n^{it} + \beta_0^i (1 - Y_n^{it})$$

$$+ \beta_1 \left(\frac{\sum_{p=1}^{t-1} I(s_n^i, s_n(t-p)) \cdot \pi_n(s_n^i, s_{-n}(t-p))}{\sum_{h=1}^{J_n} N_n^h (t-1)} \right) + \beta_2 \sum_{p=1}^{t-1} \prod_{q=1}^{p} I(s_n^i, s_n(t-q))$$

$$+ \beta_3 V_n^i (t-1)(1 - Y_n^{it}) + \varepsilon_n^{it}$$

6.3. Behavioral Reinforcement Learning Behavioral Interpretation

Table 6.3 shows the empirical results for the behavioral reinforcement case. The average payoff (normalized by each player's number of trials) and the cumulative sequence of the player's choice variables in the behavioral reinforcement learning model are statistically significant. Two variables capture the effects of a given player's learning, based on their own payoff and the bid price choice. Players with high payoffs for a certain alternative bid price throughout the game prefer to choose that bid price in order to secure more payoffs. Players who choose high bid prices are considered risk takers, since the probability of winning with the high bid is less than winning with a low bid. In addition, players having low payoffs cannot take the risk of choosing a high bid price choice as often as players with high payoffs. Such players prefer to win, even though they receive lower payoffs associated with a low bid price.

Since these players do not know their opponents' type and future payoffs, their decisions are related more to their own past payoffs and choices. Hence, the coefficient for the cumulative sequence of choice variable is positive and significant. However, the coefficient for the previous deterministic utility (habit persistence I) is negative, as in the epistemic learning model (see Table 6.1). Habit persistence represents the intensity of habit formation and introduces the concept of non-separability of preferences over time. Under this condition, an increase in current utility lowers the marginal utility in the current period and increases it in the next period. Intuitively, more players have a positive preference for that alternative in the current game, and players are more likely to choose that alternative in the next game.

Table 6.3 Estimated parameters of the behavioral reinforcement learning[20] DMNP model using behavioral reinforcement data for the two-player experiment

Attributes	Estimates	T-value
Initial Value (β_I)	0.011	11.325
Constant(alt1, β_0^1)	-0.014	-6.480
Constant(alt2, β_0^2)	-0.013	-1.848
Constant(alt3, β_0^3)	0.018	3.560
Avg. payoff (β_1)	0.008	6.871
Sequence of choice (β_2)	0.009	3.540
Habit persistence 1 (β_3)	-0.013	-2.367
ρ_1	-0.001	-1.788
ρ_2 (Habit persistence 2)	0.018	4.264
γ_{12}^{12}	0.019	6.877
γ_{12}^{13}	0.031	2.643
γ_{12}^{14}	0.002	2.690
γ_{12}^{23}	0.016	4.639
γ_{12}^{24}	0.015	2.418
γ_{12}^{34}	0.004	5.217
γ_{21}^{12}	0.011	26.486
γ_{21}^{13}	0.037	4.780
γ_{21}^{14}	0.004	4.489
γ_{21}^{23}	-0.011	-3.904
γ_{21}^{24}	-0.033	-2.305
γ_{21}^{34}	0.015	4.219
Log-likelihood at convergence	-1.4180	

---

[20] Behavioral Reinforcement Learning Model:

$$U_n^{it} = \beta_I Y_n^{it} + \beta_0^i (1 - Y_n^{it})$$

$$+\beta_1 \left(\frac{\sum_{p=1}^{t-1} I(s_n^i, s_n(t-p)) \cdot \pi_n(s_n^i, s_{-n}(t-p))}{\sum_{h=1}^{J_n} N_n^h (t-1)} \right) + \beta_2 \sum_{p=1}^{t-1} \prod_{q=1}^{p} I(s_n^i, s_n(t-q))$$

$$+\beta_3 V_n^i (t-1)(1 - Y_n^{it}) + \varepsilon_n^{it}$$

The estimates for the variance and covariance terms for bid price choice are significant at a reasonable confidence level (Tables 6.1– 6.3). This indicates that we must specify the error incorporating the impact of a player's competitive bid price choice. In particular, the coefficients for the error term for competition that impact the bid price choice for each player (γ_{nm}^{ij} and γ_{mn}^{ij}) differ, which implies that each player responds differently to other player's choice decisions, relative to the competitive impact of his or her own decisions.

6.4. Statistical Test for Epistemic vs. Behavioral Reinforcement Learning Model

We observe that some players do not follow the epistemic learning rule over all games, even though they have full information about competitors' actions. As previously mentioned, the behavioral reinforcement model estimation results were provided using the epistemic learning datasets to verify whether the player behaved in a manner consistent with the behavioral reinforcement learning process in the epistemic learning game. Thus, we applied the observed datasets from epistemic learning game to the epistemic and behavioral reinforcement model structures (see Table 6.4 for the estimation results).

Table 6.4 Estimated parameters of the epistemic (restricted) learning and behavioral reinforcement (unrestricted) learning DMNP models using the epistemic learning datasets (two-player experiment)

	Epistemic		Behavioral Reinforcement	
Attributes	Estimates	T-value	Estimates	T-value
Initial Value (β_I)	-0.842	-5.836	0.011	11.325
Constant(alt1, β_0^1)	-0.087	-8.720	-0.014	-6.480
Constant(alt2, β_0^2)	-0.712	-3.369	-0.013	-1.848
Constant(alt3, β_0^3)	-0.507	-4.412	0.018	3.560
Possible Payoff (β_1)	1.770	4.505	-	-
Avg. payoff (β_2)	-	-	0.008	6.871
Sequence of choice (β_3)	-	-	0.009	3.540
Habit persistence 1 (β_4)	-0.288	-3.296	-0.013	-2.367
ρ_1	0.408	4.823	-0.001	-1.788
ρ_2 (Habit persistence 2)	0.369	9.950	0.018	4.264
γ_{12}^{12}	0.356	2.634	0.019	6.877
γ_{12}^{13}	0.481	2.006	0.031	2.643
γ_{12}^{14}	-0.005	-3.568	0.002	2.690
γ_{12}^{23}	0.003	7.173	0.016	4.639
γ_{12}^{24}	-0.003	-1.639	0.015	2.418
γ_{12}^{34}	0.029	4.500	0.004	5.217
γ_{21}^{12}	-0.594	-1.989	0.011	26.486
γ_{21}^{13}	-0.008	-7.579	0.037	4.780
γ_{21}^{14}	0.853	1.302	0.004	4.489
γ_{21}^{23}	-1.241	-4.569	-0.011	-3.904
γ_{21}^{24}	1.162	4.144	-0.033	-2.305
γ_{21}^{34}	-0.081	-2.810	0.015	4.219
Log-likelihood at convergence	-0.4175		-0.9039	

The following procedure shows the difference between both cognitive learning models and verifies the effects of the related payoffs and sequential choice variables on bidding behavior.

In Table 6.4, both learning models show the statistical significance of the player's bidding behavior. To explain this behavior, we must determine whether the epistemic and behavioral reinforcement models differ. On one hand, only one learning model can be used to describe the player's choice behavior; on the other hand, we can conclude that both models are useful for explaining the bid price choice behavior in the epistemic learning game. The Hausman specification test is performed to determine this.

First proposed by Hausman (1978), this test evaluates model mis-specification. We first test the following null hypothesis:

H_0 : Our model is mis-specified against the alternative model

The Hausman test statistic is

$$H = [\beta_u - \beta_r]'[V_u - V_r]^{-1}[\beta_u - \beta_r]$$

where

β_u: the coefficients of variables in behavioral reinforcement learning (unrestricted) model

β_r: the coefficient of variables in the epistemic learning (restricted) model

V_u : the variance of variables in behavioral reinforcement learning (unrestricted) model

V_r: the variance of variables in epistemic learning (restricted) model

This incorporates the test conducted between the restricted model *r*, estimated for the epistemic learning models in Table 6.4, and the full model *u*, estimated for the behavioral reinforcement learning model in Table 6.4. If the model is correctly specified, then the null hypothesis will be rejected, and we can conclude that two models are different and the epistemic learning model structure is well-specified. The test statistic $H = [\beta_u - \beta_r]'[V_u - V_r]^{-1}[\beta_u - \beta_r]$ is asymptotically chi-square distributed with K_r degrees of freedom, where K_r is the number of coefficients in the restricted choice set model; β_u and β_r are the coefficient vectors estimated for the unrestricted and restricted choice sets, respectively; and V_u and V_r are the variance-covariance matrices for the unrestricted and restricted choice sets, respectively. The Hausman statistics value for the two-player experiment is:

$$H = [\beta_u - \beta_r]'[V_u - V_r]^{-1}[\beta_u - \beta_r] = 265.05$$

The chi-square value for 20 degrees of freedom and 95% confidence level is 31.41 ($\chi^2_{20,0.05} = 31.41$). Because the chi-square value is less than 265.05, we can reject this null hypothesis. We can also conclude that both models are well-specified and that there is difference between two learning models to describe player's bidding behavior.

This result demonstrates that both the epistemic and behavioral reinforcement model specifications are different but useful to describe the player's bid choice behavior in the epistemic learning game. It also implies that players behave in a manner consistent with both cognitive learning processes. This means that the mixed type of learning rule can be applied, since the player's bidding dynamics cannot be fully explained by either cognitive learning model for the two-player experiment. From this, we can suggest the general type of learning model specification to explain player's behaving in both cognitive learning rules. Thus, the mixed learning model is investigated for this general learning model to include two learning models as a special case. This mixed learning model can be better than the epistemic and behavioral reinforcement learning model to describe the player's bidding behavior by incorporating both elements from the epistemic and behavioral reinforcement learning processes.

We must specify the different model specifications and structures for the mixed learning process. The estimation procedure and model structures for mixed learning process were presented in the Chapter 4. The mixed learning model estimation results for the two-player experiment using the DMNP model are shown in the following section.

6.5. <u>DMNP Mixed Learning Models</u>

Mixed learning models integrate appropriate elements of behavioral reinforcement and epistemic learning approaches. We have already shown that both learning models can describe the player's bidding behavior for the epistemic learning game. Accordingly, the mixed learning model can be useful to improve their predictive accuracy, since it includes two learning models as special cases. This section demonstrates the important features of the mixed learning model. The maximum-likelihood parameter estimates are derived to test the empirical usefulness of mixed learning models using the two-player epistemic learning experiment datasets. The estimation process and the definition of parameters are described in Section 4.8. The following Table 6.5 illustrates the estimation results of the mixed learning DMNP model I using the epistemic learning data.

Table 6.5 Estimated parameters of the mixed learning DMNP model I using the epistemic learning datasets (two-player experiment)

Attributes	Estimates	T-value
Initial Value (β_I)	0.184	5.341
Constant(alt1, β_0^1)	0.086	3.643
Constant(alt2, β_0^2)	0.471	3.216
Constant(alt3, β_0^3)	-0.247	-2.534
Possible Payoff (β_1)	0.838	2.504
Avg. payoff (β_2)	0.467	4.591
Sequence of choice (β_3)	0.132	3.344
Habit persistence 1 (β_4)	-0.092	-6.242
α_1^E	0.344	3.846
α_2^E	0.072	4.506
ρ_1	-0.168	-3.208
ρ_2 (Habit persistence 2)	-0.577	-4.458
γ_{12}^{12}	-0.433	-1.995
γ_{12}^{13}	-0.060	-3.485
γ_{12}^{14}	-0.575	-4.670
γ_{12}^{23}	-0.062	-5.683
γ_{12}^{24}	0.078	3.641
γ_{12}^{34}	-0.602	-2.760
γ_{21}^{12}	-0.207	-4.471
γ_{21}^{13}	0.627	5.200
γ_{21}^{14}	0.001	4.804
γ_{21}^{23}	0.294	7.204
γ_{21}^{24}	0.490	3.041
γ_{21}^{34}	0.432	1.902
Log-likelihood at convergence	-0.994	

The mixed learning model I includes all explanatory variables in the epistemic and behavioral reinforcement models: initial attraction, possible payoff, average payoff, sequence of choice, and habit persistence. In this model, the players' utilities of strategies are affected by the number of choices for that strategy, the average payoffs for the strategies provided, and the expected payoffs associated with the combination of strategies among players, which are updated after each game. In addition, the mixed learning models require the specification of the player's degree of propensity for each epistemic and behavioral reinforcement learning rule (α_n^E and α_n^R).

According to Table 6.5, the coefficients of explanatory variables in mixed learning model I are statistically significant. Players who feel positive about their strategies and payoffs have a greater utility for the bid price that provided more payoffs and wins. Habit persistence has a negative coefficient; these estimation results are consistent with the results in Section 6.2 for the epistemic and behavioral reinforcement models using the epistemic learning datasets. The degree of propensity for the epistemic learning rule is 0.344 for Player 1 and 0.072 for Player 2; the degree of propensity for the behavioral reinforcement learning rule is 0.656 for Player 1 and 0.928 for Player 2. The probability of both players behaving in a manner consistent with the epistemic learning is less than the probability for the behavioral reinforcement learning players. However, Player 2 has a greater probability to be a behavioral reinforcement player over all games than Player 1.

The mixed learning model I assumes that players' choice decisions corresponding to his/her cognitive learning type are independent from other players' cognitive learning rules. This means that the competitor's cognitive learning belief on the bid price choice does not affect the player's choice of cognitive learning rules and related decisions. However, we expect that the players will dynamically adjust to the competitive environment and make the best response by observing competitors' choice decisions and their cognitive learning rules and beliefs. To evaluate the interaction among players, the estimation results for the mixed learning model II are presented in Table 6.6.

The coefficient for the explanatory variables in mixed learning model II (Table 6.6) is statistically significant, and the behavioral interpretations for the effect of those explanatory variables on players' bidding behavior are same as the mixed learning model I estimation results. As explained in Section 4.8, mixed learning model II includes parameters for the degree of propensity for mixed cognitive learning processes among players. It shows both players' preferences for each type of mixed cognitive learning. For two players, there are four types of mixed learning processes, described in Table 6.7.

In Table 6.6, the parameter values for the degree of propensity for each mixed type of learning process demonstrate that both players are more likely to behave according to the behavioral reinforcement learning process in 58% of games. This implies that players have strong interaction when both players behave in a manner consistent with the behavioral reinforcement learning process. Both players exhibit epistemic learning behavior for 25% of games; this behavior is consistent with the behavioral interpretation shown in mixed learning model I, which implies that both players' decisions are strongly involved with the behavioral reinforcement learning process.

To test the independence of players' cognitive learning processes, estimation results from both mixed learning models are provided in Tables 6.7 and 6.8.

Table 6.6 Estimated parameters of the mixed learning DMNP model II using the epistemic learning datasets (two-player experiment)

Attributes	Estimates	T-value
Initial Value (β_I)	0.849	2.377
Constant(alt1, β_0^1)	0.204	6.339
Constant(alt2, β_0^2)	0.747	7.551
Constant(alt3, β_0^3)	-0.198	-5.195
Possible Payoff (β_1)	0.301	5.063
Avg. payoff (β_2)	0.460	5.573
Sequence of choice (β_3)	0.044	1.920
Habit persistence 1 (β_4)	-0.546	-11.412
α_{12}^{EE}	0.251	7.115
α_{12}^{ER}	0.000	2.021
α_{12}^{RR}	0.581	2.861
ρ_1	0.369	21.606
ρ_2 (Habit persistence 2)	0.127	2.951
γ_{12}^{12}	-0.374	-2.200
γ_{12}^{13}	-0.106	-3.337
γ_{12}^{14}	0.313	1.821
γ_{12}^{23}	0.266	5.552
γ_{12}^{24}	0.035	2.005
γ_{12}^{34}	-0.455	-2.117
γ_{21}^{12}	-0.223	-2.516
γ_{21}^{13}	0.028	2.310
γ_{21}^{14}	-0.291	-5.783
γ_{21}^{23}	-0.667	-3.142
γ_{21}^{24}	0.783	2.506
γ_{21}^{34}	-0.050	-5.416
Log-likelihood at convergence	-1.240	

Table 6.7 The mixed cognitive learning (MCL) index and estimation results from the mixed learning DMNP model II for the two-player experiment

Player 1	Player 2		α_n^m	# of 80 games[21]
Epistemic	Epistemic	α_{12}^{EE}	0.251	20
Epistemic	Behavioral Reinforcement	α_{12}^{ER}	0.000	0
Behavioral Reinforcement	Epistemic	α_{12}^{RE}	0.168	14
Behavioral Reinforcement	Behavioral Reinforcement	α_{12}^{RR}	0.581	46

Table 6.8 The mixed cognitive learning (MCL) index and estimation results from the mixed learning DMNP model I for the two-player experiment

Player 1	Player 2		α_n^m [22]	# of 80 games
Epistemic	Epistemic	α_{12}^{EE}	0.025	2
Epistemic	Behavioral Reinforcement	α_{12}^{ER}	0.319	26
Behavioral Reinforcement	Epistemic	α_{12}^{RE}	0.047	4
Behavioral Reinforcement	Behavioral Reinforcement	α_{12}^{RR}	0.609	48

---

[21] $\text{number of game} = \alpha_n^m \times \text{total number of game}$

[22] $$\left\{ \begin{array}{l} \alpha_{12}^{EE} = \alpha_1^E \times \alpha_2^E \\ \alpha_{12}^{ER} = \alpha_1^E \times \alpha_2^R \\ \alpha_{12}^{RE} = \alpha_1^R \times \alpha_2^E \\ \alpha_{12}^{EE} = \alpha_1^R \times \alpha_2^R \end{array} \right\}$$

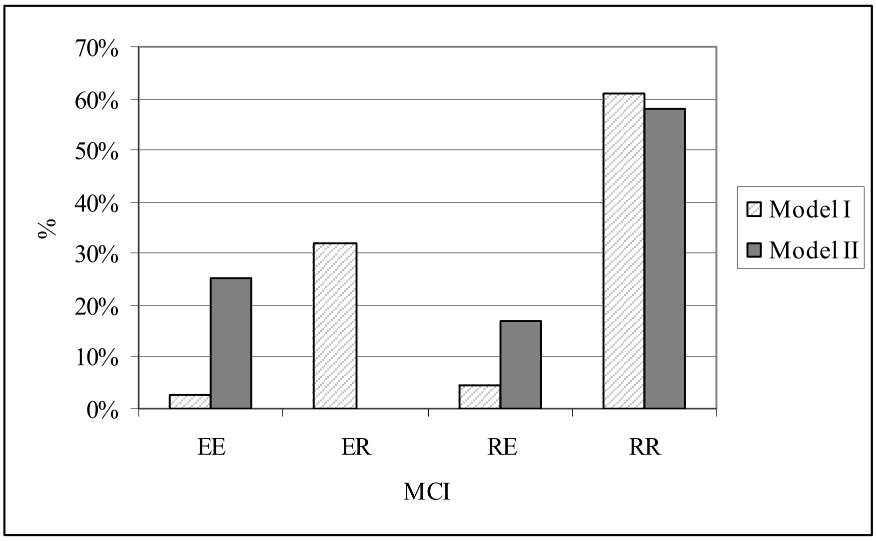


Figure 6.1 Comparison of the propensity for the mixed cognitive learning process between mixed learning DMNP models I and II

If there is no interaction among players in terms of making decisions according to their cognitive learning rules, then the degree of propensity for mixed cognitive learning processes in Table 6.7 must equal that of Table 6.8. Based on the assumption that players' cognitive learning processes are independent, the joint probability for the degree of propensity for mixed cognitive learning processes among all players (α_n^m) is calculated using the degree of propensity for each player's cognitive learning process (α_n^E and α_n^R):

$$\alpha_{12}^{EE} = \alpha_1^E \times \alpha_2^E$$

$$\alpha_{12}^{ER} = \alpha_1^E \times \alpha_2^R$$

$$\alpha_{12}^{RE} = \alpha_1^R \times \alpha_2^E$$

$$\alpha_{12}^{RR} = \alpha_1^R \times \alpha_2^R \tag{6.1}$$

Figure 6.1 compares the propensity values for the mixed cognitive learning processes among all players in mixed learning models I and II. The degree of propensity for the behavioral reinforcement learning process (RR) in both learning models is higher than any other mixed learning process. The propensity of the epistemic learning process for both players (EE) is higher than the propensity for the Player 1's epistemic and Player 2's behavioral reinforcement learning (ER) or Player 1's behavioral reinforcement and Player 2's epistemic learning (RE) in mixed learning model II. However, the propensity for the Player 1's epistemic learning and Player 2's behavioral reinforcement learning process (ER) is higher than the propensity for the epistemic learning for both players (EE) and the behavioral reinforcement learning for Player 1 and the epistemic learning for Player 2 (RE) in mixed learning model I. As shown in Figure 6.1, the propensity value for each mixed learning process differs between mixed learning model I and II, especially for the EE, ER, and RE cases. Therefore, the following statistical test is required to evaluate the difference between two mixed learning models.

Both results in Tables 6.7 and 6.8 provide the input values for the cross-tabulation and chi-square test to assess the independence of players' cognitive learning processes, which are presented in Tables 6.9 and 6.10.

Table 6.9 Model *MCL cross-tabulation for the two-player DMNP model

		Player's Mixed Cognitive Learning (MCL)				Total
		EE	ER	RE	RR	
Model I	Count	2	26	4	48	80
	% within Model I	2.5%	32.5%	5.0%	60.0%	100.0%
	% within MCL	9.1%	100.0%	22.2%	51.1%	50.0%
	% of Total	1.3%	16.3%	2.5%	30.0%	50.0%
Model II	Count	20	0	14	46	80
	% within Model I	25.0%	0.0%	17.5%	57.5%	100.0%
	% within MCL	90.9%	0.0%	77.8%	48.9%	50.0%
	% of Total	12.5%	0.0%	8.8%	28.8%	50.0%
	Count	22	26	18	94	160
	% within Model	13.8%	16.3%	11.3%	58.8%	100.0%
	% within MCL	100.0%	100.0%	100.0%	100.0%	100.0%
	% of Total	13.8%	16.3%	11.3%	58.8%	100.0%

Table 6.10 Chi-Square Tests for the two-player DMNP model

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	46.325	3	0.000
Likelihood Ratio	59.065	3	0.000
Linear-by-Linear Association	0.700	1	0.403
N of Valid Cases	160		

The chi-square test is used to find the difference between frequencies for the propensity for each mixed cognitive learning process in models I and II. The null hypothesis for the chi-square test follows:

H_0 : There is no difference between mixed learning models I and II.

In Table 6.9, the cross-tabulation result illustrates that the degree of propensity for each mixed cognitive learning process differs between the mixed learning models. In addition, the chi-square test result in Table 6.10 indicates that there is strong interaction among players' cognitive learning processes related to the choice decision. From this, we can conclude that a player considers competitors' cognitive learning type and their beliefs; hence a player switch his/her cognitive learning type according to his/her belief about compeitotor's cognitive learning type. Therefore, mixed learning model II is preferred to explain players' mixed cognitive learning behavior in sequential games.

Chapter 7. Dynamic Mixed Logit Estimation Results

7.1. Introduction

This chapter presents the empirical results and estimates of the learning models, using learning game datasets for the two and three-player sequential choice experiments. Before the estimation, we classify and specify two learning models corresponding to the two types of experimental game datasets described in Section 4.3. In addition, the mixed learning models are investigated to describe players' mixed learning behavior on bid price choice, as described in Section 4.8.

We apply the dynamic mixed logit (DML) model to existing bidding data from the two types of learning games for each experiment. Two different model structures for each cognitive learning procedure and mixed learning process are estimated to further demonstrate the capability of the developed MLE estimation procedure for DML models with a large number of parameters, alternatives, and error structures. We interpret the estimation results for learning models by evaluating the performance of the learning choice sequence. All numerical experiments in this study were implemented using FORTRAN as the programming language.

The same variance and covariance terms were assumed in v_n^{it} and Ω_n^{it} in order to simplify the estimate as follows:

$\sigma_i^2 = \sigma_j^2$, $\gamma_{nm}^{ii} = \gamma_{nm}^{jj}$, $\gamma_{nm}^{ij} = \gamma_{nm}^{ji}$, and $\gamma_{nm}^{ij} = \gamma_{nl}^{ji}$

where

i and j : alternative bid prices

n , m , and l : a players

σ_i^2 and σ_j^2 : the variance of error term

γ_{nm}^{ii} : the covariance across players for the same alternative

γ_{nm}^{ij} : the covariance across players when each player choose different choice

In our analysis, we consider several error component specifications in the DML model to introduce correlation in the utilities. The best statistical result included error components to accommodate correlation across alternatives (ρ_1), serial correlation (ρ_2), and correlation among competitors (γ_{nm}^{ij}).

As presented in Chapter 6, the DMNP estimation results using the same two-player experiment dataset were provided for the epistemic, behavioral reinforcement, and mixed learning processes. We expect that DML model will easily allow for the estimation of learning models for multiple players. Our analysis indicates that the DML structure is as statistically good as the DMNP structure.

7.2. Epistemic Learning Behavioral Interpretation

The fundamental hypothesis underlying our empirical analysis is that players' bidding decisions are largely due to sequential history of choice, payoff, and habit persistence, which are related to players' cognitive learning processes. The final variable specifications for the epistemic learning model (see Tables 7.1 and 7.2) are initial attraction, player's expected payoffs given opponents' beliefs, habit persistence, and variance-covariance terms.

Table 7.1 Estimated parameters of the epistemic learning DML model using epistemic learning data for the two-player experiment

Attributes	Estimates	T-value
Initial Value (β_I)	0.475	50.480
Constant(alt1, β_0^1)	0.118	30.481
Constant(alt2, β_0^2)	0.201	32.683
Constant(alt3, β_0^3)	-0.393	-48.633
Possible payoff (β_1)	0.707	18.520
Habit persistence 1 (β_4)	-0.188	-38.435
ρ_1	0.480	22.190
Habit persistence 2 (ρ_2)	0.130	25.713
γ_{12}^{12}	-0.060	-32.171
γ_{12}^{13}	-0.158	-190.645
γ_{12}^{14}	0.500	55.697
γ_{12}^{23}	0.098	18.685
γ_{12}^{24}	-0.033	-19.447
γ_{12}^{34}	0.241	29.687
γ_{21}^{12}	0.346	46.208
γ_{21}^{13}	-0.005	-44.532
γ_{21}^{14}	-0.026	-38.424
γ_{21}^{23}	-0.271	-29.182
γ_{21}^{24}	0.170	24.468
γ_{21}^{34}	-0.005	-110.547
Log-likelihood at convergence	-178.809	

Table 7.2 Estimated parameters of the epistemic learning DML model using epistemic learning data for the three-player experiment

Attributes	Estimates	T-value
Initial Value (β_I)	0.124	48.788
Constant(alt1, β_0^1)	0.011	18.692
Constant(alt2, β_0^2)	-0.140	-17.251
Constant(alt3, β_0^3)	-0.105	-15.528
Possible payoff (β_1)	0.392	26.599
Habit persistence 1 (β_4)	0.157	65.856
ρ_1	0.358	26.239
Habit persistence 2 (ρ_2)	0.486	27.102
γ_{1m}^{12}	-0.031	-21.051
γ_{1m}^{13}	0.035	17.095
γ_{1m}^{14}	0.280	108.171
γ_{1m}^{23}	0.139	22.320
γ_{1m}^{24}	0.070	49.181
γ_{1m}^{34}	0.006	15.357
γ_{2m}^{12}	0.088	8.172
γ_{2m}^{13}	-0.001	-9.147
γ_{2m}^{14}	-0.313	-76.756
γ_{2m}^{23}	-0.086	-26.998
γ_{2m}^{24}	0.474	79.457
γ_{2m}^{34}	-0.598	-34.679
γ_{3m}^{12}	0.003	14.364
γ_{3m}^{13}	-0.044	-65.633
γ_{3m}^{14}	0.195	20.742
γ_{3m}^{23}	-0.004	-15.754
γ_{3m}^{24}	0.107	22.955
γ_{3m}^{34}	0.063	51.468
Log-likelihood at convergence	-279.805	

The coefficients indicate the effects of these variables on players' propensity for certain bid prices. As expected, the coefficients for the expected payoffs (β_1), given a player's belief regarding their opponents' type in the epistemic learning model, are positive and strongly significant. The impact of the possible payoff[23] in Tables 7.1 and 7.2 indicates that the propensity associated with the level of bid prices increases with the possible expected payoffs, which are calculated by multiplying the probability of other players' past bid choices by the corresponding player's payoff from the mixed strategy; this is consistent with the findings presented in the previous chapter. Furthermore, players who feel more positively about the expected payoff, given a belief regarding the opponent's type, choose that bid price more often to achieve greater payoffs. Through this process, players can build beliefs about competitors' choices in the current game by observing their past choices. Given that updated belief about competitor's type, players can calculate their expected payoffs for each bid price alternative. Therefore, epistemic learning players have a greater capability for concern about their beliefs of competitors' choices than behavioral reinforcement learning players.

The results for the two-player experiment in Table 7.1 demonstrate that the utility of bid price decreases as the utility of the previous game increases, which is observed in habit persistence (β_4) term. The player behaving in a manner consistent with the epistemic learning rule is more concerned about the other player's choices and the corresponding expected payoffs than his or her own previous utility. However, in Table 7.2, the coefficient for the habit persistence for the three-player experiment is positive and significant. For the three-player experiment, since there are more possible combinations of mixed strategies among competitors, it is more difficult for

---

23 $\sum_{k=1}^{C_{-n}} \pi_n(s_n^i, s_{-n}^k(t)) \cdot p(s_{-n}^k(t) \mid s_{-n}^k(t-q), q = 1,...,t-1)$

players to predict expected payoffs. Players are required to calculate the corresponding payoffs of mixed strategies and to build their beliefs about competitors, since players have more competition. This implies that a more competitive game environment causes players to depend on their own past utility related to that bid price. Tables 7.3 and 7.4 present the DML estimation results for behavioral learning model specification using epistemic learning data.

Table 7.3 Estimated parameters of the behavioral reinforcement learning DML model using epistemic learning data for the two-player experiment

Attributes	Estimates	T-value
Initial Value (β_1)	-0.279	-29.109
Constant(alt1, β_0^1)	0.390	54.961
Constant(alt2, β_0^2)	-0.036	-79.996
Constant(alt3, β_0^3)	0.021	161.573
Avg. payoff (β_2)	0.226	92.516
Sequence of choice (β_3)	0.160	29.666
Habit persistence 1 (β_4)	-0.067	-110.777
ρ_1	0.305	33.806
Habit persistence 2 (ρ_2)	0.134	37.225
γ_{12}^{12}	0.258	47.219
γ_{12}^{13}	0.272	69.052
γ_{12}^{14}	-0.105	-89.946
γ_{12}^{23}	-0.404	-35.735
γ_{12}^{24}	0.500	215.525
γ_{12}^{34}	0.174	18.697
γ_{21}^{12}	0.334	54.440
γ_{21}^{13}	-0.056	-61.795
γ_{21}^{14}	0.005	96.417
γ_{21}^{23}	-0.161	-29.803
γ_{21}^{24}	0.037	85.682
γ_{21}^{34}	0.019	234.150
Log-likelihood at convergence	-242.526	

Table 7.4 Estimated parameters of the behavioral reinforcement learning DML model using epistemic learning data for the three-player experiment

Attributes	Estimates	T-value
Initial Value (β_1)	0.017	49.478
Constant(alt1, β_0^1)	0.157	39.994
Constant(alt2, β_0^2)	-0.647	-15.679
Constant(alt3, β_0^3)	0.223	36.748
Avg. payoff (β_2)	-0.162	-58.909
Sequence of choice (β_3)	0.125	40.044
Habit persistence 1 (β_4)	-0.100	-17.005
ρ_1	0.227	31.171
Habit persistence 2 (ρ_2)	0.711	16.612
γ_{1m}^{12}	0.057	19.962
γ_{1m}^{13}	-0.125	-13.329
γ_{1m}^{14}	-0.332	-34.825
γ_{1m}^{23}	0.178	49.494
γ_{1m}^{24}	0.283	37.323
γ_{1m}^{34}	-0.297	-35.248
γ_{2m}^{12}	-0.184	-230.516
γ_{2m}^{13}	-0.113	-42.051
γ_{2m}^{14}	-0.019	-27.437
γ_{2m}^{23}	0.329	8.154
γ_{2m}^{24}	0.094	18.358
γ_{2m}^{34}	-0.479	-23.029
γ_{3m}^{12}	-0.082	-13.943
γ_{3m}^{13}	0.375	72.809
γ_{3m}^{14}	0.500	37.009
γ_{3m}^{23}	0.351	35.394
γ_{3m}^{24}	0.181	54.912
γ_{3m}^{34}	-0.122	-89.938
Log-likelihood at convergence	-309.497	

The coefficients of explanatory variables and variance-covariance terms are statistically significant for both experiments. The signs of the coefficients of explanatory variables in the DML estimated results for the two-player experiment are same as those results in DMNP results for the two-player experiment (see Table 6.1). Therefore, the behavioral interpretation is as same as in Section 6.2. The goodness-of-fit is measured by the likelihood ratio. The log likelihood ratio value in Table 7.3 (-242.526) is less than the log likelihood ratio value in Table 7.1 (-178.809). This means that the epistemic learning model can describe players' bidding behavior better than the behavioral learning model using epistemic learning game data.

Table 7.4 presents the estimation results of the behavioral reinforcement learning model for the three-player experiment using epistemic learning data. The coefficient of explanatory variables and variance-covariance terms are statistically significant in Table 7.4 (three-player experiment). However, the coefficients for the average payoff and habit persistence I variables are negative; this differs from the estimation results for the two-player experiment in Table 7.3, implying that players consider their possible payoffs given beliefs about opponents' type rather than their own past payoffs. In other words, players are more likely to use information about opponents' past choices to predict their next action. In addition, the log likelihood ratio in Table 7.4 (-309.497) is less than the log likelihood ration in Table 7.2 (-279.805). From these results, players' bidding behavior for the three-player experiment is well explained by the epistemic learning model specification compared to the behavioral reinforcement leaning model specification for the epistemic learning game.

7.3. ***Behavioral Reinforcement Learning Behavioral Interpretation***

In this section, we apply the behavioral reinforcement learning model to the existing behavioral reinforcement learning datasets. Average payoff, cumulative sequence of choice, and habit persistence are used as the explanatory variables in the behavioral reinforcement learning model specifications in Tables 7.5 and 7.6 to illustrate the impact of these variables on players' bidding behavior in the behavioral reinforcement learning game.

Tables 7.5 and 7.6 present the estimated parameters of the explanatory variables for the behavioral reinforcement model and the error components. The average payoff[24] (normalized by each player's number of trials) and the cumulative sequence of the players' choice[25] variables in the behavioral reinforcement learning model for both experiments are positive and statistically significant. These results are consistent with the DMNP estimation results from the behavioral reinforcement learning model for the two-player experiment. The estimation results from the average payoff and sequence of choice variables show the player's behavioral learning belief about their own payoff for that bid price choice. Players believe that the bid price providing high past payoffs can continuously produce high payoffs in future games. They also believe that competitors have the same belief about their decisions, which is

---

24 $\left(\dfrac{\sum_{p=1}^{t-1} I(s_n^i(t-p)) \cdot \pi_n(s_n^i(t-p))}{\sum_{h=1}^{J_n} N_n^h(t-1)} \right)$

25 $\sum_{p=1}^{t-1} \prod_{q=1}^{p} I(s_n^i(t-q))$

reinforced by the previous preference for that bid price. Accordingly, the coefficient for the cumulative sequence of choice variable is positive and significant.

The coefficient of habit persistence in deterministic utility is negative and statistically significant (Table 7.5). As previously mentioned habit persistence represents the intensity of habit formation and introduces non-separability of preferences over time. If players do not have the opportunity to choose a certain bid price in the past or have less expected payoff for that bid price, even though they believe that it is possible to have more payoffs by choosing that bid price, they eager to choose that bid price in the future games.

In Table 7.6, the coefficient of habit persistence for the three-player experiment is positive and statistically significant. This indicates that the past choice experience for that bid price continues to provide high utility in the current period, but it makes the current choice less desirable. In the three-player experiment, players must predict more combinations of mixed strategies, because they are competing against more players. Therefore, players are not confident that they will win the next game, because there is more pressure as a result of observing competitors' reactions over games.

Table 7.5 Estimated parameters of the behavioral reinforcement learning DML model using behavioral reinforcement data for the two-player experiment

Attributes	Estimates	T-value
Initial Value (β_1)	0.313	47.337
Constant(alt1, β_0^1)	-0.003	-95.871
Constant(alt2, β_0^2)	-0.088	-17.764
Constant(alt3, β_0^3)	0.299	5.859
Avg. payoff (β_2)	0.437	31.916
Sequence of choice (β_3)	0.043	5.792
Habit persistence 1 (β_4)	-0.345	-11.629
ρ_1	0.473	34.536
Habit persistence 2 (ρ_2)	0.526	62.494
γ_{12}^{12}	-0.422	-47.573
γ_{12}^{13}	-0.140	-21.602
γ_{12}^{14}	-0.197	-25.089
γ_{12}^{23}	-0.365	-53.131
γ_{12}^{24}	-0.061	-18.170
γ_{12}^{34}	0.150	29.485
γ_{21}^{12}	0.278	6.044
γ_{21}^{13}	0.455	14.268
γ_{21}^{14}	-0.115	-37.822
γ_{21}^{23}	-0.162	-11.798
γ_{21}^{24}	0.262	17.913
γ_{21}^{34}	0.200	12.732
Log-likelihood at convergence	192.182	

Table 7.6 Estimated parameters of the behavioral reinforcement learning DML model using behavioral reinforcement data for the three-player experiment

Attributes	Estimates	T-value
Initial Value (β_1)	-0.426	-32.416
Constant(alt1, β_0^1)	0.923	61.830
Constant(alt2, β_0^2)	0.287	38.732
Constant(alt3, β_0^3)	-0.078	-31.497
Avg. payoff (β_2)	0.189	41.618
Sequence of choice (β_3)	0.363	20.317
Habit persistence 1 (β_4)	0.526	31.579
ρ_1	0.684	67.794
Habit persistence 2 (ρ_2)	0.290	26.628
γ_{1m}^{12}	-0.221	-30.369
γ_{1m}^{13}	-0.095	-92.164
γ_{1m}^{14}	-0.281	-41.969
γ_{1m}^{23}	1.012	53.951
γ_{1m}^{24}	0.253	27.814
γ_{1m}^{34}	0.092	66.266
γ_{2m}^{12}	-0.507	-37.935
γ_{2m}^{13}	0.187	15.031
γ_{2m}^{14}	0.132	50.752
γ_{2m}^{23}	-0.271	-34.383
γ_{2m}^{24}	-0.059	-83.839
γ_{2m}^{34}	-0.280	-45.416
γ_{3m}^{12}	0.454	17.938
γ_{3m}^{13}	0.384	32.052
γ_{3m}^{14}	-0.092	-61.246
γ_{3m}^{23}	0.247	39.482
γ_{3m}^{24}	-0.207	-25.815
γ_{3m}^{34}	-0.023	-47.433
Log-likelihood at convergence	-313.129	

The error components introduced in the utility function generate covariance in unobserved factors across alternatives, time, and players. The estimates for the variance and covariance terms for bid price choice are significant at a reasonable confidence level (Tables 7.1–7.6). We expect that these error terms can capture the unobserved competition effect on bidding behavior in learning games; the estimation results demonstrate that the error term incorporating the impact of a player's competition must be specified. This implies that player's decision is affected by other players' choices and corresponding expected payoffs or their own average payoffs. This is also observed in the error terms, which indicates the player's unobserved impact of competition among players on the strategy choice decision.

7.4. Statistical Test for Epistemic vs. Behavioral Reinforcement Learning Model

In Section 7.2, the estimation results show that both the epistemic and behavioral reinforcement models are well-fitted to describe the player's bidding behavior for the epistemic game. Here, we evaluate whether those learning models are well-specified. Table 7.7 shows the estimation results of both learning models for the two-player experiment using the same epistemic learning data.

Table 7.7 Estimated parameters of the epistemic (restricted) learning and behavioral reinforcement (unrestricted) DML model using epistemic learning data (two-player experiment)

	Epistemic		Behavioral Reinforcement	
Attributes	Estimates	T-value	Estimates	T-value
Initial Value (β_I)	0.475	50.480	-0.279	-29.109
Constant(alt1, β_0^1)	0.118	30.481	0.390	54.961
Constant(alt2, β_0^2)	0.201	32.683	-0.036	-79.996
Constant(alt3, β_0^3)	-0.393	-48.633	0.021	161.573
Possible Payoff (β_1)	0.707	18.520	-	-
Avg. payoff (β_2)	-	-	0.226	92.516
Sequence of choice (β_3)	-	-	0.160	29.666
Habit persistence 1 (β_4)	-0.188	-38.435	-0.067	-110.777
ρ_1	0.480	22.190	0.305	33.806
ρ_2 (Habit persistence 2)	0.130	25.713	0.134	37.225
γ_{12}^{12}	-0.060	-32.171	0.258	47.219
γ_{12}^{13}	-0.158	-190.645	0.272	69.052
γ_{12}^{14}	0.500	55.697	-0.105	-89.946
γ_{12}^{23}	0.098	18.685	-0.404	-35.735
γ_{12}^{24}	-0.033	-19.447	0.500	215.525
γ_{12}^{34}	0.241	29.687	0.174	18.697
γ_{21}^{12}	0.346	46.208	0.334	54.440
γ_{21}^{13}	-0.005	-44.532	-0.056	-61.795
γ_{21}^{14}	-0.026	-38.424	0.005	96.417
γ_{21}^{23}	-0.271	-29.182	-0.161	-29.803
γ_{21}^{24}	0.170	24.468	0.037	85.682
γ_{21}^{34}	-0.005	-110.547	0.019	234.150
Log-likelihood at Convergence	-178.809		-242.526	

To verify that the epistemic learning model is correctly specified, the Hausman specification test is applied. This test can be used to obtain an estimator that is efficient and consistent under the following hypothesis.

H_0 : Our model is mis-specified against the alternative model

The Hausman test statistic is

$$H = [\beta_u - \beta_r]' [V_u - V_r]^{-1} [\beta_u - \beta_r]$$

where

β_u : the coefficients of variables in behavioral reinforcement learning (unrestricted) model

β_r : the coefficients of variables in epistemic learning (restricted) model

V_u : the variances of variables in behavioral reinforcement learning (unrestricted) model

V_r : the variances of variables in epistemic learning (restricted) model

The Hausman statistics value is:

$$H = [\beta_u - \beta_r]' [V_u - V_r]^{-1} [\beta_u - \beta_r] = 432764.2$$

The Hausman statistical value for the two learning models is 432764.2. The chi-square value for 20 degrees of freedom and 95% confidence level is 31.41 ($\chi^2_{20,0.05} = 31.41$). This Hausman statistical value (432764.2) is higher than the chi-square value (31.41); therefore, the null hypothesis is rejected. This result is consistent with the findings in Section 6.4. The Hausman test evaluates the significance of estimators compared to an alternative estimator. Therefore, this result demonstrates that the epistemic learning model is well-specified and that there is a difference between the learning models in describing bidding behavior.

Table 7.8 provides the estimation results of the epistemic and behavioral reinforcement learning models for the three-player experiment.

Table 7.8 Estimated parameters of the epistemic (restricted) learning and behavioral reinforcement (unrestricted) learning DML model using epistemic learning data (three-player experiment)

	Epistemic		Behavioral Reinforcement	
Attributes	Estimates	T-value	Estimates	T-value
Initial Value (β_I)	0.124	48.788	0.017	49.478
Constant(alt1, β_0^1)	0.011	18.692	0.157	39.994
Constant(alt2, β_0^2)	-0.140	-17.251	-0.647	-15.679
Constant(alt3, β_0^3)	-0.105	-15.528	0.223	36.748
Possible Payoff (β_1)	0.392	26.599	-	-
Avg. payoff (β_2)	-	-	-0.162	-58.909
Sequence of choice (β_3)	-	-	0.125	40.044
Habit persistence 1 (β_4)	0.157	65.856	-0.100	-17.005
ρ_1	0.358	26.239	0.227	31.171
ρ_2 (Habit persistence 2)	0.486	27.102	0.711	16.612
γ_{1m}^{12}	-0.031	-21.051	0.057	19.962
γ_{1m}^{13}	0.035	17.095	-0.125	-13.329
γ_{1m}^{14}	0.280	108.171	-0.332	-34.825
γ_{1m}^{23}	0.139	22.320	0.178	49.494
γ_{1m}^{24}	0.070	49.181	0.283	37.323
γ_{1m}^{34}	0.006	15.357	-0.297	-35.248
γ_{2m}^{12}	0.088	8.172	-0.184	-230.516
γ_{2m}^{13}	-0.001	-9.147	-0.113	-42.051
γ_{2m}^{14}	-0.313	-76.756	-0.019	-27.437
γ_{2m}^{23}	-0.086	-26.998	0.329	8.154
γ_{2m}^{24}	0.474	79.457	0.094	18.358
γ_{2m}^{34}	-0.598	-34.679	-0.479	-23.029
γ_{3m}^{12}	0.003	14.364	-0.082	-13.943
γ_{3m}^{13}	-0.044	-65.633	0.375	72.809
γ_{3m}^{14}	0.195	20.742	0.500	37.009
γ_{3m}^{23}	-0.004	-15.754	0.351	35.394
γ_{3m}^{24}	0.107	22.955	0.181	54.912
γ_{3m}^{34}	0.063	51.468	-0.122	-89.938
Log-likelihood at convergence	-279.805		-309.497	

The Hausman specification test is applied to evaluate the difference between the epistemic and behavioral reinforcement learning models. The Hausman statistic can be viewed as a measure of distance between the epistemic and the behavioral reinforcement learning model estimators. If this distance is short, then these models are not significantly different. The null hypothesis is:

H_0: Our model is mis-specified against the alternative model

The Hausman statistic value for two learning models is:

$$H = [\beta_u - \beta_r]'[V_u - V_r]^{-1}[\beta_u - \beta_r] = 91739.61$$

The chi-square value for 26 degrees of freedom and 95% confidence level is 38.89 ($\chi^2_{21,0.05} = 38.89$). The above Hausman static value (91739.6) is much higher than the chi-square value (38.89). From this chi-square test result, the above null hypothesis can be rejected. Accordingly, we conclude that the epistemic learning model for the three-player experiment is well-specified and that there is a difference between the epistemic and behavioral reinforcement learning model specifications in the epistemic learning game. Thus, both learning models describe the player's bidding behavior well but in different ways.

These estimation results and Hausman statistical tests show that players did not behaved consistent according to only one of the cognitive learning rules in sequential games. The results of the Hausman test prove that both models are well-specified and fit the epistemic learning game datasets well. Based on this, we cannot determine which model is better to describe behavior in the epistemic learning game. Therefore, the mixed learning model should be developed to describe this player's mixed cognitive bidding behavior. This mixed learning model can describe the learning process and can capture the impact of players' interactive cognitive learning processes. The next section presents the mixed learning models.

7.5. <u>DML Mixed Learning Models</u>

We found that some players did not use all of the available information, such as competitors' past choices in the epistemic game. Accordingly, some decisions were made by players based on the behavioral reinforcement learning rule, even though they participating in the epistemic learning game. The either or both of cognitive learning rules determined how players behaved during each round of the epistemic learning game. The estimation results in the previous section support this phenomenon. Based on this, we propose model specifications and estimation procedures for the mixed learning process described in Chapter 4.

The dynamic mixed logit model structure can provide the estimation results for two types of mixed learning models. The first model assumes that players behave independently according to either cognitive learning rule for each game and do not consider the competitors' cognitive learning rules before making a decision in the epistemic learning game. For the second model, the players' cognitive learning beliefs are dynamically affected by competitors' cognitive learning rules, maximizing their utility in the epistemic learning game. Therefore, mixed learning model II can capture the effect of interdependence among players for cognitive learning process on bidding behavior. Tables 7.9 and 7.10 show the estimation results for the mixed learning model I for the two-player experiment using epistemic learning data.

Table 7.9 Estimated parameters of the mixed learning DML model I using epistemic learning data (two-player experiment)

Attributes	Estimates	T-value
Initial Value (β_I)	-0.100	-26.618
Constant(alt1, β_0^1)	0.058	28.599
Constant(alt2, β_0^2)	0.023	9.601
Constant(alt3, β_0^3)	-0.087	-35.257
Possible Payoff (β_1)	0.002	34.981
Avg. payoff (β_2)	0.322	6.847
Sequence of choice (β_3)	0.250	13.824
Habit persistence 1 (β_4)	-0.056	-28.276
α_1^E	0.441	10.464
α_2^E	0.194	15.621
ρ_1	0.613	38.410
ρ_2 (Habit persistence 2)	0.375	85.485
γ_{12}^{12}	0.346	33.498
γ_{12}^{13}	-0.246	-33.748
γ_{12}^{14}	-0.118	-29.336
γ_{12}^{23}	-0.075	-52.230
γ_{12}^{24}	0.023	37.169
γ_{12}^{34}	0.099	76.517
γ_{21}^{12}	-0.202	-12.283
γ_{21}^{13}	0.580	56.308
γ_{21}^{14}	0.245	43.217
γ_{21}^{23}	-0.156	-42.791
γ_{21}^{24}	-0.116	-31.736
γ_{21}^{34}	-0.165	-17.361
Log-likelihood at convergence	-177.745	

The explanatory variables in mixed learning model I have a significant positive impact on bid choice behavior. The coefficients and behavioral interpretation for those results are consistent with the findings presented in Section 7.2. The coefficients for the expected payoffs and choice experience related to the player's epistemic and behavioral reinforcement learning process are positive and strongly significant. The degree of propensity (α_n^E) for players 1 and 2 in the behavioral reinforcement learning process is almost 0.6 and 0.8, respectively, over 80 games. The players' bidding preference increases with the increasing effects of all explanatory variables included in the behavioral reinforcement learning model.

Mixed learning model I assumes that players independently follow cognitive learning rules without considering competitors' cognitive rules. Table 7.9 indicates that the degree of propensity for the epistemic learning rule for players 1 and 2 is 0.441 and 0.194, respectively. This implies that both players prefer to conduct themselves like the behavioral reinforcement learning players. Player 2 is more likely to behave according to the behavioral reinforcement learning process than Player 1. Player 1 is more likely to behave according to the epistemic learning process than Player2. These results imply that, for the two-player experiment, players apply both cognitive learning rules in the decision-making process for all games.

Table 7.10 Estimated parameters of the mixed learning DML model II using epistemic learning data (two-player experiment)

Attributes	Estimates	T-value
Initial Value (β_I)	-0.551	-41.718
Constant(alt1, β_0^1)	-0.306	-12.438
Constant(alt2, β_0^2)	-0.107	-13.877
Constant(alt3, β_0^3)	-0.134	-22.762
Possible Payoff (β_1)	0.299	52.079
Avg. payoff (β_2)	0.283	20.734
Sequence of choice (β_3)	0.201	34.548
Habit persistence 1 (β_4)	-0.011	-14.332
α_{12}^{EE}	0.170	18.665
α_{12}^{ER}	0.166	108.246
α_{12}^{RR}	0.502	18.533
ρ_1	0.509	31.540
ρ_2 (Habit persistence 2)	0.128	19.782
γ_{12}^{12}	0.123	18.898
γ_{12}^{13}	0.301	26.945
γ_{12}^{14}	-0.394	-78.561
γ_{12}^{23}	-0.009	-52.424
γ_{12}^{24}	0.443	128.395
γ_{12}^{34}	0.050	39.778
γ_{21}^{12}	0.232	40.825
γ_{21}^{13}	-0.435	-30.693
γ_{21}^{14}	0.453	31.839
γ_{21}^{23}	0.086	138.850
γ_{21}^{24}	0.500	17.756
γ_{21}^{34}	0.026	45.805
Log-likelihood at convergence	-201.622	

Table 7.10 illustrates the estimation results of mixed learning model II for the two-player experiment using the epistemic learning game datasets. This model II can capture the effect of the interaction among players' cognitive learning processes on the bidding behavior. This impact is represented by the parameters (α_n^m) as the

degree of propensity for the m type of mixed cognitive learning process over all games, which is also presented in Table 7.11 for the two-player experiment. The coefficients for possible payoff, average payoff, and sequence of choice are positive, indicating that players are more likely to choose that bid price with high expected payoffs in the epistemic learning model, high average payoffs, and more recent choice experience in the behavioral reinforcement learning model. The coefficient for habit persistence is negative, which is consistent with the results in Section 7.2.

Table 7.11 The mixed cognitive learning (MCL) index and estimation results from the mixed learning DML model II for the two-player experiment

Player 1	Player 2		α_n^m [26]	# of observation
Epistemic	Epistemic	α_{12}^{EE}	0.170	14
Epistemic	Behavioral Reinforcement	α_{12}^{ER}	0.166	13
Behavioral Reinforcement	Epistemic	α_{12}^{RE}	0.162	13
Behavioral Reinforcement	Behavioral Reinforcement	α_{12}^{RR}	0.502	40

Table 7.11 presents the mixed cognitive learning index and estimation results associated with MCL from mixed learning model II. The parameter for the behavioral

26
$$\begin{cases} \alpha_{12}^{EE} = \alpha_1^E \times \alpha_2^E \\ \alpha_{12}^{ER} = \alpha_1^E \times \alpha_2^R \\ \alpha_{12}^{RE} = \alpha_1^R \times \alpha_2^E \\ \alpha_{12}^{RR} = \alpha_1^R \times \alpha_2^R \end{cases}$$

reinforcement learning propensity for both players (α_{12}^{RR}) is the highest value for the mixed cognitive learning processes. This is consistent with the findings for the two-player DMNP mixed learning model II shown in Table 6.6. We expect this phenomenon, since both players also show a high degree of propensity for the behavioral reinforcement learning process in the DML mixed learning model I (see Table 7.10). However, this does not mean that both models describe players' mixed learning processes in the same way. Table 7.12 presents the mixed cognitive learning index and the estimation results associated with MCL from mixed learning model I; we could compare these results to the previous estimation results in Table 7.11.

Table 7.12 The mixed cognitive learning (MCL) index and estimation results from the mixed learning DML model I for the two-player experiment

Player 1	Player 2		α_n^m	# of observation
Epistemic	Epistemic	α_{12}^{EE}	0.086	7
Epistemic	Behavioral Reinforcement	α_{12}^{ER}	0.355	28
Behavioral Reinforcement	Epistemic	α_{12}^{RE}	0.108	9
Behavioral Reinforcement	Behavioral Reinforcement	α_{12}^{RR}	0.451	36

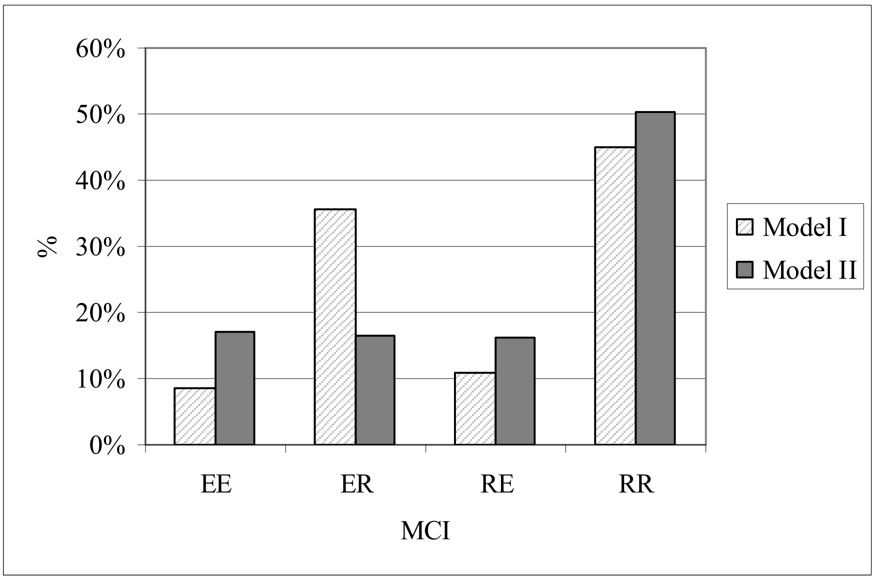


Figure 7.1 Comparison of the propensity for the mixed cognitive learning process between mixed learning DML models I and II

The joint probability equation (6.1) in Section 6.5 provides the parameter values for each type of mixed learning process among players, based on that assumption that players' cognitive learning processes are dependent on one another. In Figure 7.1, both models show the players' high propensity for the behavioral reinforcement learning process (RR). In mixed learning model II, the propensity for three types of mixed learning processes among players (EE, ER, and RE) is nearly identical. In mixed learning model I, Player 1's propensity for epistemic learning and Player 2's propensity for behavioral reinforcement learning (ER) is higher than the other two types of mixed learning processes (EE and RE), which demonstrates the difference between the two modeling frameworks. The following cross-tabulation and chi-square test are applied to determine the difference between the models in terms of describing mixed cognitive learning behavior.

Table 7.13 Model *MCL cross-tabulation for the two-player DML model

		Player's Mixed Cognitive Learning (MCL)				Total
		EE	ER	RE	RR	
Model I	Count	7	28	9	36	80
	% within Model I	8.8%	35.0%	11.3%	45.0%	100.0%
	% within MCL	33.3%	68.3%	40.9%	47.4%	50.0%
	% of Total	4.4%	17.5%	5.6%	22.5%	50.0%
Model II	Count	14	13	13	40	80
	% within Model I	17.5%	16.3%	16.3%	50.0%	100.0%
	% within MCL	66.7%	31.7%	59.1%	52.6%	50.0%
	% of Total	8.8%	8.1%	8.1%	25.0%	50.0%
	Count	21	41	22	76	160
	% within Model	13.1%	25.6%	13.8%	47.5%	100.0%
	% within MCL	100.0%	100.0%	100.0%	100.0%	100.0%
	% of Total	13.1%	25.6%	13.8%	47.5%	100.0%

Table 7.14 Chi-Square Tests for the two-player DML model

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	8.759	3	0.033
Likelihood Ratio	8.938	3	0.030
Linear-by-Linear Association	0.124	1	0.725
N of Valid Cases	160		

The chi-square statistical test results in Table 7.14 show that the mixed learning models describe players' behavior differently. More specifically, player's bidding behavior is affected by other competitors' epistemic or behavioral cognitive learning process. This result is consistent with the results in Section 6.5 for the two-player mixed learning DMNP models.

As noted, the additional datasets for the different auction types and multiple-player cases were collected to test the dynamic strategy choice modeling framework. The basic estimation results for each cognitive learning model are presented in Tables 7.1 and 7.2. Here, two additional mixed learning models are investigated using the three-player experimental datasets. Tables 7.15 and 7.16 present the estimation results.

Table 7.15 Estimated parameters of the mixed learning DML model I using epistemic learning data (three-player experiment)

Attributes	Estimates	T-value
Initial Value (β_I)	0.330	151.824
Constant(alt1, β_0^1)	-0.217	-37.420
Constant(alt2, β_0^2)	-0.334	-86.953
Constant(alt3, β_0^3)	-0.292	-65.758
Possible Payoff (β_1)	0.182	40.314
Avg. payoff (β_2)	0.004	192.846
Sequence of choice (β_3)	0.555	47.691
Habit persistence 1 (β_4)	0.129	152.221
α_1^E	0.480	24.535
α_2^E	0.335	42.865
α_3^E	0.801	156.102
ρ_1	0.343	47.004
ρ_2 (Habit persistence 2)	0.363	31.535
γ_{1m}^{12}	-0.174	-262.313
γ_{1m}^{13}	0.019	27.345
γ_{1m}^{14}	0.707	49.719
γ_{1m}^{23}	-0.231	-79.349
γ_{1m}^{24}	-0.091	-52.204
γ_{1m}^{34}	-0.280	-94.530
γ_{2m}^{12}	0.072	38.562
γ_{2m}^{13}	0.312	41.965
γ_{2m}^{14}	-0.206	-25.220
γ_{2m}^{23}	0.584	54.203
γ_{2m}^{24}	0.440	32.980
γ_{2m}^{34}	-0.433	-152.919
γ_{3m}^{12}	-0.486	-41.135
γ_{3m}^{13}	0.276	29.024
γ_{3m}^{14}	0.388	44.104
γ_{3m}^{23}	-0.260	-136.209
γ_{3m}^{24}	-0.115	-47.564
γ_{3m}^{34}	0.491	136.097
Log-likelihood at convergence	-344.833	

The coefficients for the explanatory variables included in Table 7.15 are positive and statistically significant. This implies that past payoffs and choice experience are important determinants of bidding strategy. The degree of propensity for the epistemic learning process for Players 1, 2, and 3 is 0.480, 0.335, and 0.801, respectively. The propensity for the behavioral reinforcement learning propensity is 0.520, 0.665, and 0.199 for Players 1, 2, and 3, respectively. Mixed learning model I assumes that there is no interaction among players in terms of making decisions according to their cognitive learning rules; hence, a player only relies on their own cognitive learning rules to make their decision, but they can switch learning rules by considering payoffs and choice experience. Under this assumption, Players 1 and 2 behave according to the behavioral reinforcement learning process in over 50% of all games, while Player 3 behaves according to the epistemic learning rule for almost 80% of all epistemic games.

Table 7.16 presents the mixed learning model II estimation results for the three-player experiment. The statistical significance and coefficient signs for the explanatory variables are same as those for the estimation results in Section 7.2. Accordingly, the behavioral interpretations related to those variables remain consistent. The coefficient for average payoff is negative, which is consistent with the results presented in Table 7.4, since mixed learning model II combines elements from both the epistemic and behavioral reinforcements learning models.

Table 7.16 Estimated parameters of the mixed learning DML model II using epistemic learning data (three-player experiment)

Attributes	Estimates	T-value
Initial Value (β_I)	0.0143	56.893
Constant(alt1, β_0^1)	-0.0095	-39.357
Constant(alt2, β_0^2)	-0.0150	-102.142
Constant(alt3, β_0^3)	-0.0133	-98.177
Possible Payoff (β_1)	0.0080	100.789
Avg. payoff (β_2)	-0.0008	-200.474
Sequence of choice (β_3)	0.0249	36.455
Habit persistence 1 (β_4)	0.0068	56.804
α_{123}^{EEE}	0.0220	59.790
α_{123}^{EER}	0.0942	208.301
α_{123}^{ERE}	0.0647	140.693
α_{123}^{ERR}	0.1664	86.668
α_{123}^{REE}	0.0391	63.507
α_{123}^{RER}	0.0000	-99.299
α_{123}^{RRR}	0.0000	-76.174
ρ_1	0.0153	25.221
ρ_2 (Habit persistence 2)	0.0160	468.396
γ_{1m}^{12}	-0.0072	-83.772
γ_{1m}^{13}	0.0019	80.596
γ_{1m}^{14}	0.0319	57.231
γ_{1m}^{23}	-0.0100	-81.452
γ_{1m}^{24}	-0.0045	-55.092
γ_{1m}^{34}	-0.0123	-49.235
γ_{2m}^{12}	0.0037	171.160
γ_{2m}^{13}	0.0138	63.165
γ_{2m}^{14}	-0.0086	-143.835
γ_{2m}^{23}	0.0264	80.806
γ_{2m}^{24}	0.0198	29.456
γ_{2m}^{34}	-0.0199	-46.507
γ_{3m}^{12}	-0.0217	-43.350
γ_{3m}^{13}	0.0123	35.141
γ_{3m}^{14}	0.0175	43.303

γ_{3m}^{23}	-0.0122	-65.604
γ_{3m}^{24}	-0.0064	-72.919
γ_{3m}^{34}	0.2771	30.589
Log-likelihood at convergence	-423.292	

Table 7.17 shows the index and estimation results for the propensity of m type of mixed cognitive learning process from the mixed learning DML model II. The degree of propensity coefficients for seven types of mixed cognitive learning is statistically significant, as shown in Table 7.16.

Table 7.17 The mixed cognitive learning (MCL) index and estimation results from the mixed learning DML model II for the three-player experiment

Player 1	Player 2	Player 3		α_n^m	# of observation
Epistemic	Epistemic	Epistemic	α_{123}^{EEE}	0.022	2
Epistemic	Epistemic	Behavioral Reinforcement	α_{123}^{EER}	0.094	8
Epistemic	Behavioral Reinforcement	Epistemic	α_{123}^{ERE}	0.065	5
Epistemic	Behavioral Reinforcement	Behavioral Reinforcement	α_{123}^{ERR}	0.166	13
Behavioral Reinforcement	Epistemic	Epistemic	α_{123}^{REE}	0.039	3
Behavioral Reinforcement	Epistemic	Behavioral Reinforcement	α_{123}^{RER}	0.000	0
Behavioral Reinforcement	Behavioral Reinforcement	Behavioral Reinforcement	α_{123}^{RRR}	0.000	0
Behavioral Reinforcement	Behavioral Reinforcement	Epistemic	α_{123}^{RRE}	0.614	49

In mixed learning model II, each player's decision is affected by the competitor's cognitive learning rules. Therefore, the interaction among players in terms of making decisions according to their cognitive learning rules result in the above parameters for the degree of propensity for each type of mixed cognitive learning process.

The estimation results in Table 7.17 indicate that the degree of propensity for Players 1 and 2 in the behavioral reinforcement learning process and Player 3 in the epistemic learning process (α_{123}^{RRE}) is 0.614, which is the highest value for all mixed cognitive learning processes. This phenomenon is consistent with the results presented in Table 7.15 (α_1^E=0.480, α_2^E=0.335, and α_3^E=0.801). Players 1 and 2 behave more like behavioral reinforcement learning players, while Player 3 behaves more like the epistemic learning player. However, from this result, we cannot conclude that both mixed learning models are the same statistically, which means that there is independence among players in terms of making decisions according to their cognitive learning rules.

The following chi-square test for independence is used to evaluate the statistically significant difference between proportions for each *m* type of degree of propensity for players' mixed cognitive learning processes from the mixed learning models, using epistemic learning dataset. To perform the chi-square test, we recall the estimation results in Table 7.15. From these results (α_1^E=0.480, α_2^E=0.335, and

α_3^E =0.801), we provide the joint probability[27] values for each mixed type of cognitive learning process (α_n^m). In mixed learning model I, we assume the independence of cognitive learning processes among players. Therefore, the joint probability of mixed three-player cognitive learning process is calculated by:

$$\alpha_{123}^{EEE} = \alpha_1^E \times \alpha_2^E \times \alpha_3^E,\ \alpha_{123}^{EER} = \alpha_1^E \times \alpha_2^E \times \alpha_3^R,$$

$$\alpha_{123}^{ERE} = \alpha_1^E \times \alpha_2^R \times \alpha_3^E,\ \alpha_{123}^{ERR} = \alpha_1^E \times \alpha_2^R \times \alpha_3^R,$$

$$\alpha_{123}^{REE} = \alpha_1^R \times \alpha_2^E \times \alpha_3^E,\ \alpha_{123}^{RER} = \alpha_1^R \times \alpha_2^E \times \alpha_3^R,$$

$$\alpha_{123}^{RRR} = \alpha_1^R \times \alpha_2^R \times \alpha_3^R,\ \alpha_{123}^{RRE} = \alpha_1^R \times \alpha_2^R \times \alpha_3^E \quad (7.1)$$

Table 7.18 shows the joint probability for the each type of cognitive learning process, calculated using the mixed learning model I estimation results for each player's degree of propensity for each type of cognitive learning process (α_1^E =0.480, α_2^E =0.335, and α_3^E =0.801). Using these values along with the estimation results in Table 7.17, we provide the following cross-tabulation analysis and chi-square results in Tables 7.19 and 7.20 for the three-player experiment.

[27] Joint probability is the probability of two events in conjunction and the probability of both events together.

Table 7.18 The mixed cognitive learning (MCL) index and estimation results from the mixed learning DML model I for three-player experiment

Player 1	Player 2	Player 3		α_n^m [28]	# of observation
Epistemic	Epistemic	Epistemic	α_{123}^{EEE}	0.129	10
Epistemic	Epistemic	Behavioral Reinforcement	α_{123}^{EER}	0.032	3
Epistemic	Behavioral Reinforcement	Epistemic	α_{123}^{ERE}	0.256	20
Epistemic	Behavioral Reinforcement	Behavioral Reinforcement	α_{123}^{ERR}	0.064	5
Behavioral Reinforcement	Epistemic	Epistemic	α_{123}^{REE}	0.140	11
Behavioral Reinforcement	Epistemic	Behavioral Reinforcement	α_{123}^{RER}	0.035	3
Behavioral Reinforcement	Behavioral Reinforcement	Behavioral Reinforcement	α_{123}^{RRR}	0.069	6
Behavioral Reinforcement	Behavioral Reinforcement	Epistemic	α_{123}^{RRE}	0.277	22

---

[28] $$\begin{Bmatrix} \alpha_{123}^{EEE} = \alpha_1^E \times \alpha_2^E \times \alpha_3^E \\ \alpha_{123}^{EER} = \alpha_1^E \times \alpha_2^E \times \alpha_3^R \\ \alpha_{123}^{ERE} = \alpha_1^E \times \alpha_2^R \times \alpha_3^E \\ \alpha_{123}^{ERR} = \alpha_1^E \times \alpha_2^R \times \alpha_3^R \end{Bmatrix}, \begin{Bmatrix} \alpha_{123}^{REE} = \alpha_1^R \times \alpha_2^E \times \alpha_3^E \\ \alpha_{123}^{RER} = \alpha_1^R \times \alpha_2^E \times \alpha_3^R \\ \alpha_{123}^{RRR} = \alpha_1^R \times \alpha_2^R \times \alpha_3^R \\ \alpha_{123}^{RRE} = \alpha_1^R \times \alpha_2^R \times \alpha_3^E \end{Bmatrix}$$

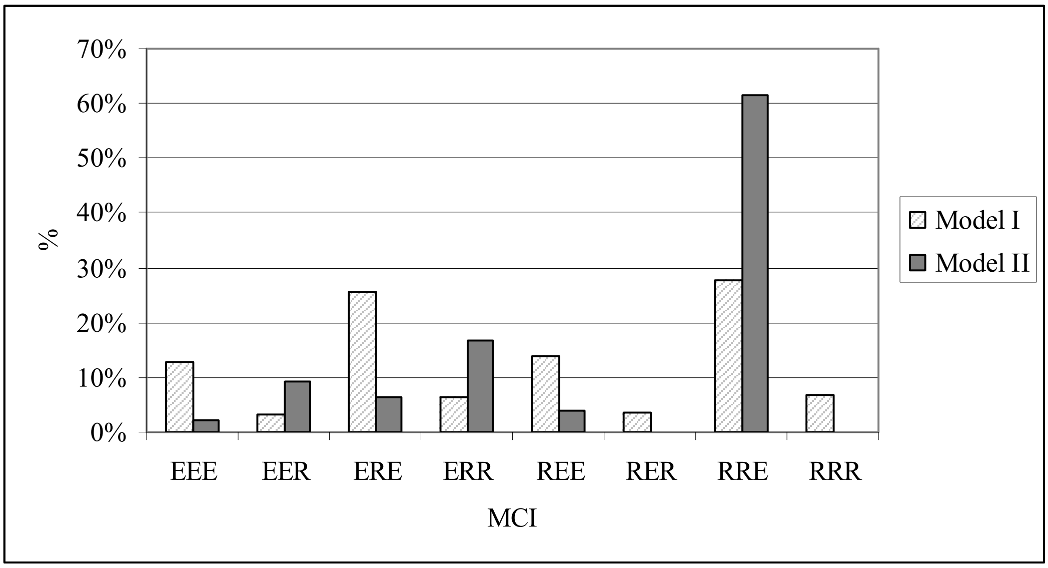


Figure 7.2 Comparison of the propensity for the mixed cognitive learning process between mixed learning DML models I and model II (three-player experiment)

Figure 7.2 presents the propensity among players to behave according to each mixed cognitive learning process. There are eight types of mixed learning processes for the three-player experiment. The propensity for the RRE[29] type is higher than for any other mixed cognitive learning process for both models. However, the degree of propensity for the RRE mixed learning process provided by mixed learning model II is much higher than mixed learning model I estimation results. Mixed learning model I provides a higher value of propensity for the ERE mixed learning process compared to the other types (EEE, EER, ERR, REE, RER, and RRR).

The following cross-tabulation displays the joint distribution of the estimation results

---

[29] Players 1 and 2 in the behavioral reinforcement process and Player 3 in the epistemic learning process

from the mixed learning models I and II. It shows the different degrees of propensity for each type of mixed cognitive learning process among players. To evaluate the significant difference between the models, the chi-square test is performed and the test results are included in Table 7.20. The null hypothesis is that there is no interaction among players in terms of making decisions according to their cognitive learning rules.

Table 7.20 demonstrates that we can reject that null hypothesis, which means that player's cognitive learning belief on the choice decision is correlated to competitor's choice decision associated with his/her cognitive learning process. Therefore, the mixed learning model II is preferred to mixed learning model I.

Table 7.19 Model *MCL cross-tabulation for the three-player DML model

		Player's Mixed Cognitive Learning (MCL)								
		EEE	EER	ERE	ERR	REE	RER	RRR	RRE	Total
Model I	Count	10	3	20	5	11	3	6	22	80
	% within Model I	12.5%	3.8%	25.0%	6.3%	13.8%	3.8%	7.5%	27.5%	100.0%
	% within MCL	83.3%	27.3%	80.0%	27.8%	78.6%	100.0%	100.0%	31.0%	50.0%
	% of Total	6.3%	1.9%	12.5%	3.1%	6.9%	1.9%	3.8%	13.8%	50.0%
Model II	Count	2	8	5	13	3	0	0	49	80
	% within Model II	2.5%	10.0%	6.3%	16.3%	3.8%	0%	0%	61.3%	100.0%
	% within MCL	16.7%	72.7%	20.0%	72.2%	21.4%	0%	0%	69.0%	50.0%
	% of Total	1.3%	5.0%	3.1%	8.1%	1.9%	0%	0%	30.6%	50.0%
Total	Count	12	11	25	18	14	3	6	71	160
	% within Model	7.5%	6.9%	15.6%	11.3%	8.8%	1.9%	3.8%	44.4%	100.0%
	% within MCL	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
	% of Total	7.5%	6.9%	15.6%	11.3%	8.8%	1.9%	3.8%	44.4%	100.0%

Table 7.20 Chi-Square Tests for the three-player DML model

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	44.001	7	.000
Likelihood Ratio	49.368	7	.000
Linear-by-Linear Association	10.575	1	.001
N of Valid Cases	160		

Chapter 8. Conclusions

8.1. Summary of Research Accomplishments

The main objective of this dissertation is to develop a theoretical framework and methodology to model a player's dynamic strategy choice behavior under the competitive environment, by extending and adapting the existing model framework. This research is different from the previous studies in experimental economics and traditional transportation demand studies. First of all, this dissertation introduces the advanced discrete choice model, and cognitive learning process into the modeling of bidder's behavior, in contrast to the traditional game theory approach. In particular, the multinomial probit and mixed logit modeling frameworks were applied to the estimation of dynamic strategy choice behavior models, for the data obtained from two experiments, by allowing for a more flexible correlation structure of the error term. Furthermore, we present the concept of competition impact into dynamic choice model structures. Therefore, this dissertation introduces a new perspective on transportation behavior modeling in competitive environments to show how learning models can describe a player's choice behavior in a sequential auction type of game. Additionally, it shows the possibility of how those learning processes can be combined to make a better model and combines methodological strengths of earlier studies by estimating the cognitive learning models.

The interaction among players' decisions was examined with the observations of bidding behavior, since the decision problem is related to players' strategies due to the interdependence of competitors' bids, costs, and payoffs. Two experiments were conducted in which decision-makers are in hypothetical bidding situations. The first experiments included two players who each separately participated 80 times in the two types of games, and who were able to choose between four discrete alternative bid prices each time, and the first price auction game was applied. The second experiment included three players with the same conditions as the first experiment except the auction type; the second price auction was applied. Initially, this dissertation focuses on the two types of cognitive learning processes; thus two types of games were performed related to two types of cognitive learning rules: epistemic and behavioral reinforcement learning processes. The behavioral reinforcement game provided subjects with enough information about the player's own payoffs and his/her choice records, while the epistemic learning game gave full information about the player's own experience, as well as competitors' choice history. Those experiments were performed to observe actual players' game-to-game dynamic decision in a different type of auction game. The data acquired from the two experiments allows researchers to investigate the processes determining players' strategies, choice decisions in response to different levels of cognitive capacity and information.

In the previous works from the experimental economics, the simple power or logit form of probability choice modeling framework was applied to experimental datasets which included the dynamic bidding choice behavior. In this dissertation, the dynamic multinomial probit and dynamic mixed logit models were used to allow more flexible model specification through parameters in the variance-covariance matrix. Moreover, the unobserved influences of a player's competitive interaction on the decision-making process, state dependence, and auto correlation were captured in error structures.

Two learning model structures were developed for the strategy choice in a sequential auction type of game addressing both epistemic and behavioral reinforcement learning processes. The epistemic learning model assumes that in a repeated game, players always know their own characteristics, but they have incomplete information regarding their opponents' learning belief, or rationality. Players can draw the distribution of competitors' action by observing the whole sequence of the others' decisions. A player does not know the opponents' cognitive learning types, or belief from their choice decisions, but they can conjecture the competitors' instrumental rationality, which is some form of stationarity of the competitors' behavior (Walliser, 1998). In the behavioral reinforcement learning model, players with limited rationality have knowledge about their own opportunities and payoffs in repeated auction games, but they could not predict the type of opponent or his/her past actions and payoffs. Players revised their experience according to the utility obtained from each strategy in past actions. In this process, the player imitates past actions reinforced by successful strategies in past games. Hence, a player's utility is updated by observing his or her sequence of actions and corresponding payoffs.

The estimation results from two cognitive learning models were provided to describe the players' bidding behavior. It confirmed that players who hold to the epistemic learning rule are more likely to use all information about their opponent's type to make a decision. Furthermore, in the epistemic learning game, players also adopted the behavioral reinforcement learning rule, as well as the epistemic learning rule in strategy decisions. The Hausman statistical test was performed to test for model miss-specification. The test results illustrated that players did not always behave in a consistent manner with only one of the cognitive learning rules in sequential games. Through the Hausman test, the epistemic learning model was found to be well specified and fitted to the epistemic learning game datasets. In the behavioral reinforcement game, the estimation results demonstrated that players relied on their own historical payoffs, and choice decisions to determine their strategy for the next game because that player does not know the opponent's type and future payoffs.

In the epistemic learning game, the estimation results showed that both epistemic and behavioral reinforcement models were well fitted to describe the players' bidding behavior for the epistemic game. The results implied that players behave in a manner with the mixed learning process, which incorporates both elements from the epistemic and behavioral reinforcement learning processes. Players can switch the strategic learning belief from either one of the two learning processes to the other (epistemic and behavioral reinforcement). Therefore, the mixed learning model was investigated to explain the players' bidding behavior corresponding to the mixed learning process.

Two types of mixed learning models were investigated in a sequential epistemic learning auction game. The mixed learning models incorporated appropriate elements of behavioral reinforcement and epistemic learning approaches and were found to be useful in improving their predictive accuracy, since it included two learning models as special cases. The models also utilized all the information provided by the players and the game environment. One mixed learning model made an assumption about the interdependence among players' cognitive learning processes, while the other showed the independence of these processes. Both mixed learning models provided the players' degree of propensity toward each type of mixed learning process. We compared and examined these mixed learning structures for parameters with clear psychological interpretations using statistical chi-square tests. These tests proved that that player's cognitive learning belief on the choice decision was correlated to competitor's choice decision associated with his/her cognitive learning process.

The error components provide the covariance in unobserved factors across alternatives, times, and players. The estimates for the variance and covariance terms for bid price choice were statistically significant at a reasonable confidence level. From The coefficients for the error term for competition that impact the bid price choice for each player were different, which implies that each player responds differently to other players' choice decisions, relative to the competitive impact of his or her own decisions. Consequently, we found that the error term incorporating the impact of a player's competition on bid price choice needs to be specified. This demonstrates how important it is for the model to capture the effect of unobserved interactions between competitors. By transforming an unobservable into an observable interaction, we can witness directly how parameter estimates change when new information is introduced

The perspective adopted in this study is that of an analyst or observer (which may also be a competitor) seeking to predict the outcome of the bidding process followed by a player in a repeated auction game. The dynamic strategy choice model framework was presented to recognize, in specification and parameterization, the nature of multi-player interaction and its dynamics over multiple bids (plays). A dynamic strategy choice model, which uses concepts from both discrete choice theory and game theory to model bidder (carrier) behavior in repeated auction-type games, can explain people's adjustment patterns when dealing with complicated transportation-related decision-making problems. This type of modeling can be applied in spot market situations related to transportation auctions.

8.2. Applications and Future Researches

While motivated by and intended for freight service procurement marketplaces, the model framework and structure developed in this study is applicable to other competitive situations that entail repeated decisions over time. For example, air carriers must continually make service and pricing decisions in anticipation of or in response to competitors' actions. Urban travelers have cognitive learning rules that correspond to behavioral reinforcement learning in a low information environment. They do not model the impact of their actions on other passengers' utilities or their environment. However, using information technology in an urban network system allows people to behave according to epistemic learning rules; they consider other traveler's reactions to the IT system with respect to route choice and mode choice.

For the freight service market, the dynamic choice model can be applied to carriers' bidding behavior in an Internet transportation marketplace. Carriers model the behavior of other carriers and shippers in order to maximize profits. Given the assumption of the auction and the player types of a strategic problem, carriers may behave according to epistemic or behavioral reinforcement learning rules, depending upon their level of rationality. Shippers also predict behavior using the dynamic strategy choice model; thus, their actions can be predicted depending on their rational and computational capabilities and the complexity of their own logistic problems. In addition, air travel agents can use this model framework to consider interactions between carriers and passengers. Based on this information, they can model their own behavior to obtain profits.

There are many ways to extend this study for future research. This study introduced two characteristic types of cognitive learning processes. Two separate model structures were presented in the methodology section, using different assumptions of players' skill to collect information on player interactions. Both models can be applied in the low information environment. With the behavioral reinforcement learning model, if a player does not know his or her own payoff, a different model specification is required. The epistemic model only uses the form of stationarity of opponent type (past choices). In epistemic learning model, we must be careful to specify the type of beliefs that must be used as inputs into these models. The model can be improved by collecting more information on such factors as opponent payoffs. The use of this information allows for the player's behavior to be explained in certain types of games. However, the above model specification, incorporating low information, is important in the construction of a more general model in real low information environments.

From the myopic perspective on the learning model specification, we do not consider the depreciation effect of the amount of game experience used for calculating the average payoff, cumulative sequence of choice, and probability of competitor's choice decision variable values. As players have more experience, the effect of the previous experience on the current choice will be depreciated. Therefore, we need to consider those parameters for each variable included in model specifications for both learning models.

We used two experimental datasets for two and three-player cases. As a result, the models are limited to applying to various types of auction game datasets. Therefore, we need more experiments to provide the general model specification for each cognitive learning process. Moreover, if we can have more players, the model structures for the error terms and utility specification can be changed and the subsequent estimation results will be different. We expect that a more competitive environment allows players to behave in different ways to show the adjustment learning patterns compared to a less competitive environment. Furthermore, we need to consider the impact on the players' bidding behavior between on-line and off-line sequential games. The number of players in a game is fixed throughout all games in this dissertation. However, the number of bidders in online auctions is likely to vary for each game. Therefore, more uncertainty can exist in on-line game than in non-online auctions. Those uncertainty factors should be modeled and included as the component of the utility function, and as unobserved noise factors.

Despite the above contribution and possible applications, we do not believe that our learning model captures all significant aspects of learning in games. In this dissertation, we did not consider the private information for each player, which can determine their cost, and more public information (e.g. fleet size, number of players, auction type, player's previous experience, fleet status at a given time; location or time window of available truck on schedule, etc). We need to specify what other factors can affect the decision-making process in the game, and how we can model it differently. Therefore, more specific public or private information parameters can be considered in the utility function, in order to specify more unobserved factors.

Bibliography

Anderson, S.P., J.K. Goeree., and C.A. Holt (1999). The logit equilibrium: a perspective on intuitive behavioral anomalies, working paper, University of Virginia.

Anderson S. P., J.K. Goeree, and C.A. Holt (2001). Minimum-effort coordination games: stochastic potential and logit equilibrium. *Games and Economic Behavior* 34, 177-199.

Arora, A., H. Xu, R. Padman, and W. Vogt (2003). Optimal Bidding in Sequential Online Auctions, working paper, The H.John Heinz III School of Public Policy Management, Carnegie Mellon University.

Arthur, B. (1991). Designing economic agents that act like human agents: A behavioral approach to bounded rationality, AER Proceedings 81, 353-359.

Basu, K. (1994). The Traveler's Dilemma: Paradoxes of Rationality in Game Theory, *American Economic Review*, 84(2), 391-395.

Ben-Akiva, M., and S. R. Lerman (1987). *Discrete Choice Model: Theory and Application to Travel Demand.* MIT Press. Cambridge. Mass.

Ben-Akiva, M., D. Bolduc, and J. Walker (2001). Specification, Identification, & Estimation of the Logit Kernel (or Continuous Mixed Logit) Model. Working paper, MIT, 2001.

Binmore, K. (1987). Modeling rational players I, *Economics and Philosophy*, 3, 9-55.

Bichler, M. (2001). *The Future of e-Markets: Multidimensional Market Mechanism,* Cambridge University Press, 2001.

Bhat, C. R. (1995). A Heteroscedastic Extreme Value Model of Intercity Travel Mode Choice. *Transportation Research Part B*, Vol. 29, No. 6, 471-483.

Borgers, T., and R. Sarin (2000). Naïve Reinforcement Learning with Endogenous Aspirataion, International Economic Review, Vol. 41, No. 4, 921-950

Boylan, R., and M. El-Gamal (1993). Fictitious Play: A Statistical Study of Multiple Economic Experiments, Games and Economic Behavior, 5, 205–222.

Braun, P.A., G.M. Constantidines, and W.E. Ferson (1993). Time Nonseparability in Aggregate Consumption, *European Economic Review*, 1993, *37* (5), 897-920.

Bunch, D. (1991). Estimability in the Multinomial Probit Model, *Transportation Research B*, 25B, No. 1, 1-12.

Camerer, C.F., and C. Anderson (2000). Experience-Weighted Attraction Learning in Sender-Receiver Signaling Games, *Economic Theory*, 16, 689-718.

Camerer, C.F., and T-H. Ho (1998). EWA Learning in Coordination Games: Probability Rules, Heterogeneity, and Time Variation, *Journal of Mathematical Psychology* 42, 305-326.

Camerer, C.F., and T-H. Ho (1999). Experience-Weighted Attraction Learning in Normal Form Games, *Econometrica* 67, 837-874.

Camerer, C.F., T-H. Ho, and J.K. Chong (2002). Sophisticated EWA Learning and Strategic Teaching in Repeated Games, *Journal of Economic Theory*, 104:1, 137-88.

Capra, C., J.K. Monica, R.G. Goeree, and C.A. Holt (1997). Anomalous Behavior in a Traveler's Dilemma?, Working Paper, University of Virginia.

Chen, H-C, J.W. Friedman, and J.C. Thisse (1997). Boundedly Rational Nash Equilibrium: A Probabilistic Choice Approach, *Games and Economic Behavior* 18 (1), 32-54.

Cheung, Y. W., and D. Friedman (1997). Individual Learning in Normal Form Games: Some Laboratory Results, Games and Economic Behavior, 19, 46–76.

Dufwenberg, M., and U. Gneezy (2000). Price Competition and Market Concentration: An Experimental Study, *International Journal of Industrial Organization* 18, 7-22.

Feltovich, N. (2000). Reinforcement-Based vs. Belief-Based Learning Models in Experimental Asymmetric-Information Games, Econometrica, 68, 605–641.

Figliozzi, M.A. (2004). Performance and Analysis of Spot Truck-Load Procurement Markets Using Sequential Auctions, PhD Dissertation, University of Maryland.

Figliozzi, M.A., H.S. Mahmassani, and P. Jaillet (2003a). A Framework for the study of Carrier Strategies in an Auction Based Transportation Marketplace, *Transportation Research Record* 1854, 162-170.

Figliozzi, M.A., H.S. Mahmassani, and P. Jaillet (2003b). Modeling Carrier Behavior in Sequential Auction Transportation Markets, *10th International Association of Travel Behaviour Research Conference*, Lucerne, Switzerland, July 2003.

Figliozzi, M.A., H.S. Mahmassani, and P. Jaillet (2005). Impacts of Auction Settings on the Performance of Truckload Transportation Marketplaces, *Transportation Research Record* 1906, 89-96.

Fudenberg, D., and D. Levine (1998). *Theory of Learning in Games*, MIT Press. Cambridge. Mass.

Hajivassiliou, V. A., D. McFadden, and P. Ruud. (1996). Simulation of multivariate normal rectangle probabilities and their derivatives: Theoretical and computational results. *J. Econometrics* 72, 85–134.

Geweke, J, M. Keane, and D. Runkle (1994). Alternative Computational Approaches to Inference in the Multinomial Probit Model, *Review of Economics and Statistics*, 76, 609-632.

Goeree, J.K., and C.A. Holt (1999). Stochastic game theory: for playing games, not just for doing theory, proceeding of the National Academy of Science, 96, 10564-10567.

Goyal, S., and M. Janssen (1995). Can we rationally learn to coordinate? *Theory and Decision*, 40 (1), 29-49.

Hausman, J. (1978). Specification tests in econometrics, *Econometrica* 46(6), 1251-71.

Jofre-Bonet, M., and M. Pesendorfer (2003). Estimation of a Dynamic Auction Game, *Econometrica*, Vol. 71, No 5, 1443-1489

Jou, R.C., and H.S. Mahmassani (1994). Day-toDay Dynamics of Commuter Travel Behavior in an Urban environment: Departure time and Route Decisions.

Proceedings for the 6th International Conference on Travel Behavior, IATBR, Vol.1, Santiago, Chile, 210-222

Jou, R.C. (1994). A Model of Dynamci Commuter Behavior Incorporating Trip-Chaining, PhD Dissertation, University of Texas at Austin

Kagel, J. (1995). Auctions: A Survey of Experimental Research. In Handbook of Experimental Economics, edited by J. Kagel and A. Roth. Princeton, NJ: Princeton University Press, 501–585.

Lam, S.H., and H.S. Mahmassani (1991). Multinomial probit estimation: Computational procedures and applications. Methods for Understanding Travel Behavior in the 1990's. Proceeding for the 6th International Conference on Travel Behavior, IATBR, Vol.1, Quebec, Canada, 228-242.

Lam, S.H. (1991). Multinomial Probit Model Estimation: Computational Procedures and Applications, PhD Dissertation, University of Texas at Austin.

Lee, J.S., and B.K. Szymanski (2005). A Novel Auction Mechanism for Selling Time-Sensitive E-Services Proceedings of the 7th IEEE International Conference on E-Commerce Technology, Vol. 0, 75-82

Lerman, S., and C. Manski (1981). On the Use of Simulated Frequencies to Approximate Choice Probabilities. In *Structural Analysis of Discrete Data with Econometric Applications*, Charles Manski and Daniel McFadden, eds. Cambridge, MA, MIT Press.

Liu, Y.H. (1997). Dynamics of Commuter Behaviour under Real-Time Traffic Information, PhD Dissertation, University of Texas at Austin.

Manski, C., and D. McFadden (1981). *Structural Analysis of Discrete Data With Econometric Applications*, MIT Press

Mahmassani, H.S., and R.C. Jou (1996). Bounded rationality in commuter decision dynamics: incorporating trip chaining in departure time and route switching decisions. Presented at The Conference on Theoretical Foundations of Travel Choice Modelling, Stockholm, Sweden.

McAfee, R. P., and P.J. McMillan (1987a). Auctions and Bidding, Journal of Economic Literature, Vol.25, 699-738.

McAfee, R., and P.J. McMillan (1997). Auction and Bidding, *Journal of Economic Literature*, Vol. 25, 699 – 738.

McFadden, D. (1981). Econometric Models of Probabilistic Choice. In *Structural Analysis of Discrete Data with Econometric Applications*, Charles Manski and Daniel McFadden, eds. Cambridge, MA: MIT Press.

McFadden, D. (1989). A Method of Simulated Moments for Estimation of Discrete Response Models without Numerical Integration. *Econometrica* 57(5), 995-1026.

McFadden, D. and K. Train (2000). Mixed MNL Models for Discrete Response, *Journal of Applied Econometrics,* 15(5), 447-470.

Milgrom, P. (1998). Game Theory and the Spectrum Auctions, *European Economic Reviews*, 42, 771-778

Mookherjee, D., and B. Sopher (1994). Learning Behavior in Experimental Matching Pennies, Games and Economic Behavior, 7, 62–91.

Mookherjee, D., and B. Sopher (1997). Learning and Decision Costs in Experimental Constant-Sum Games, Games and Economic Behavior, 19, 97–132.

Nyarko, Y., and A. Schotter (2002). An experimental Study of Belief Learning Using Elicited Beliefs, Econometrica, Vol 70, No3, 971-1005

Osborne, M. (2004). *Introduction to Game Theory.* Oxford University Press.

Rankin, F., J. Van Huyck, and R. Battalio (1997). Strategic Similarity and Emergent Conventions: Evidence From Scrambled Payoff Perturbed Stag-Hunt Games, Mimeo, Department of Economics, Texas A&M University.

Roth, A., and I. Erev (1995). Learning in Extensive form Games: Experimental Data and Simple dynamic Models in the Intermediate Run, Games and Economic Behavior, 6, 164-212

Roth, A., and I. Erev (1998). Predicting How People Play Games: Reinforcement Learning in Experimental Games With Unique, Mixed Strategy, Equilibria, American Economic Review, 88, 848–881.

Satin, R., and F. Vahid (1999). Predicting How People Play Games: A Simple Dynamic Model of Choice, working paper, Monash University.

Srinivasan, K.K. (2000). Dynamic Decision and Adjustment Processes in Commuter Behavior under Real-Time Traffic Information, PhD Dissertation, University of Texas at Austin.

Srinivasan, K. and H. Mahmassani (2003). Analyzing Heterogeneity and Unobserved Structural Effects in Route-Switching Behavior Under ATIS: a Dynamic Kernel Logit Formulation, *Transportation Research B* 37(9),793-814.

Srinivasan, K.K., and H.S. Mahmassani (2005). A Dynamic Kernel Logit Model for the Analysis of Longitudinal Discrete Choice Data: Properties and Computational Assessment, Transportation Science, Vol. 39, No. 2, 160–181

Train, K. (1986). Qualitative Choice Analysis, MIT Press.

Train, K. (2003). *Discrete Choice Methods with Simulation.* Cambridge University Press.

Walliser, B. (1998). A spectrum of equilibration processes in game theory, *Journal of Evolutionary Economics* 8, 67-87.

Weibull. J.W. (1995). *Evolutionary Game Theory*. MIT press.