Title
Modeling Customer Behavior in Loyalty Programs

Author
Taylor, Wayne

Publication Date
2017

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

Modeling Customer Behavior in Loyalty Programs

A dissertation submitted in partial satisfaction

of the requirements for the degree

Doctor of Philosophy in Management

by

Wayne Taylor

2017

ABSTRACT OF THE DISSERTATION

Modeling Customer Behavior in Loyalty Programs

by

Wayne Taylor

Doctor of Philosophy in Management

University of California, Los Angeles, 2017

Professor Anand V. Bodapati, Chair

Loyalty programs have exploded in popularity in recent decades. In the United States alone, membership has reached 1.3 billion (Ferguson and Hlavinka, 2007). In spite of their continued popularity, the effectiveness of these programs has been long debated in the literature, with mostly mixed results. Verhoef (2003) finds that the effects are positive but very small, DeWulf et al. (2001) finds no support for positive effects of direct mail, Shugan (2005) finds that firms gain short term revenue at the expense of longer term reward payments, and Hartmann and Viard (2008) found no evidence of the loyalty program creating switching costs. Rather than attempt to broadly label loyalty programs as either effective or ineffective, this dissertation instead focuses on how firms can use their loyalty program databases to model customer behavior. In the first essay I investigate how positive and negative casino experiences influence the casino's targeting strategy. In the second essay I study a coalition loyalty program and use the variation in

the coalition network size to estimate the value of store participation. Finally, I extend the first two essays and summarize the ongoing debate as to whether the human brain processes information using Bayesian inference. The intent of this research is to both contribute to academic literature and also provide insight for practitioners to improve decision making.

In the first essay I consider the direct marketing targeting problem in situations where 1) the customer's experience quality level varies from occasion to occasion, 2) the firm has measures of these quality levels, and 3) the firm can customize marketing according to these measures and the customer's behaviors. A primary contribution of this paper is a framework and methodology that allows the manager to assess the marketing response of a forward-looking customer with any specific experience and behavior history, which in turn can be used to decide which customers to target for marketing. This research develops a novel, tractable way to estimate and introduce flexible heterogeneity distributions into Bayesian learning models with forward looking agents. The model is estimated using data from the casino industry, an industry which generates more than $60 billion in U.S. revenues but has surprisingly little academic, econometric research. The counterfactuals offer interesting findings on gambler learning and direct marketing responsiveness and suggest that casino profitability can increase substantially when marketing incorporates gamblers' beliefs and past outcome sequences into the targeting decision.

In the second essay I consider the problem faced by managers of coalition loyalty programs of store network composition. In a coalition loyalty program, managers need to determine which stores to include in the coalition network. The value of a particular store may depend on both the changes in projected member spend at the focal store and the changes in spend at the other stores in the network. The primary contribution of this research is a model that can measure these same-store and cross-store effects. These cross effects can be used to determine the extent of spillover both *to* and *from* a focal store from other stores in the network. I estimate the model using data from a coalition network in Europe and find that there are substantial cross-store effects. The findings have substantial implications for managers: even if an individual store is not contributing much revenue to the coalition its presence in the network can positively influence the network as a whole substantially.

In the extension I summarize the recent debate as to whether or not the human brain processes information in a Bayesian fashion and propose potential departures from rational Bayesian inference. I then present models of these departures and review the characteristics of the data sources that would be ideal in estimating these effects. This extension builds on the first two essays by 1) exploring more deeply the learning model presented in the first essay and 2) presenting models that may account for the potential imperfect recall and incomplete knowledge that was assumed in the second essay. The goal of the extension is to highlight how a clear understanding of how the brain operates and processes new experiences could have drastic implications for a loyalty program's design and targeting marketing strategy.

The dissertation of Wayne Taylor is approved.

Randolph Bucklin

Dominique Hanssens

Brett Hollenbeck

Sivaramakrishnan Siddarth

Anand V. Bodapati, Committee Chair

University of California, Los Angeles

2017

Contents

List of Figures

List of Tables

ACKNOWLEDGMENTS

I am very grateful to my doctoral advisor, Professor Anand Bodapati, for his unwavering commitment to my success. Professor Bodapati helped me even prior to my application into the doctoral program and his mentorship has persisted ever since. Over our numerous meetings together I quickly discovered his immeasurable knowledge in a wide range of areas – these interactions will be my most memorable moments from the program. The kindness emanating from Professor Bodapati is a gift and I am incredibly fortunate to have worked under him.

I want to acknowledge Professor Peter Rossi for his generosity and patience as he dedicated much time and effort to ensure my understanding of Bayesian statistics. I also appreciate the time he devoted to learning about my research, providing valuable feedback, and pushing me to go further and further. His expansive knowledge and passion for research is truly inspiring.

Professor Randolph Bucklin was instrumental in my success. He always found time to meet on short notice and his feedback on my drafts and presentations has been incredibly helpful. I was always impressed by his ability to quickly hone in on the key managerial contribution of whatever research project we were discussing. Whenever I left his office my passion for research was revitalized.

I want to also acknowledge Professor Dominique Hanssens. His ability to take on a managerial perspective has been invaluable on numerous occasions and played a significant role in shaping my research projects. I am very fortunate to have been able to work with Professor Hanssens during my time in the program.

Professor Brett Hollenbeck was a crucial resource for me on structural modeling. His mentorship was indispensable in the formation of my first essay. He provided excellent feedback and suggestions for my other research projects as well. Professor Hollenbeck was always eager to receive updates on my projects and his insights have been immensely useful.

I want to acknowledge Professor Sivaramakrishnan Siddarth for his helpful and timely feedback during my draft development and presentations. I greatly appreciate the multiple trips he made to UCLA to listen to my research updates.

I want to acknowledge Professor Don Morrison for his advice and generous support. I am also grateful to Professors Sanjog Misra, Raphael Thomadsen, Robert Zeithammer, Hal Hershfield, Aimee Drolet, Elisabeth Honka, Stephen Spiller, Daniel Oppenheimer, Suzanne Shu, Sanjay Sood, Sylvia Hristakeva, Shi Zhang, and Andres Terech for their support, feedback, and for sharing their knowledge.

I want to acknowledge the administrative support of Lydia Heyman, April Barfield, Christian Yirgu, and Craig Jessen. They were always helpful for scheduling rooms, finding supplies, or effortlessly handling any one of the potential challenges faced in the day-to-day operations of managing doctoral students.

I am very grateful to Paul Hoban and Elizabeth Webb for their mentorship and emotional support that helped me through this arduous journey.

I also must acknowledge my fellow doctoral students in marketing: Keunwoo Kim, Charlene Chu, Li Jiang, Mirei Takashima, Marissa Sharif, Daniel Yavorsky, Jonathan Lim, Kate Christensen, and Kalyan Rallabandi for their support and encouragement.

I would be remiss if I did not acknowledge Professor Sherri Kimes, who is my inspiration for deciding to pursue my doctoral degree many years ago.

Above all, I am grateful for the monumental efforts, sacrifices, and support of my family. My mother and father have always selflessly encouraged my pursuits even at the expense of the distance and time I have spent away from them. Ida and Ray helped in so many ways with moving into, around, and out of Los Angeles and by providing extra assistance when needed. Lola and Wilson provided companionship and much needed emotional support through the most trying times of the program. Finally, I am forever grateful for the support of Katrina. She has made tremendous sacrifices and has had to endure an almost equivalent amount of stress but at the same time forgo the upside of receiving a doctoral degree. Her strength and compassion is beyond words.

Thank you all very much.

VITA

2005-2009 B.Sc. in Hotel Administration

Cornell University, Ithaca, New York

2009-2012 Senior Manager of Casino Marketing Analytics

Las Vegas Sands, Las Vegas, Nevada

Invited Presentations

09/2016 University of Georgia

Athens, Georgia

10/2016 Southern Methodist University

Dallas, Texas

10/2016 Santa Clara University

San Jose, California

10/2016 University of Texas at Dallas

Richardson, Texas

04/2017 University of California, San Diego

San Diego, California

1 Incorporating Experience Quality Data into CRM Models: The Impact of Gambler Outcomes on Casino Return Times

1.1 Introduction

This paper focuses on the direct marketing problem of whom to target and with what offers in situations where the customer learns about the firm through multiple interactions. The situations I focus on have the following four specific characteristics: (a) The customer's experience quality level is random in that it varies from occasion to occasion according to a certain distribution. (b) The customer uses the experiences to learn and form beliefs about some key characteristics of the distribution. These beliefs affect the customer's expected utility from future interactions with the firm, which in turn affect future decisions on whether to interact with the firm. (c) The firm has access to measures of the customer's experience quality levels. (d) The firm can use information on these measures to make different offers to different customers. These four characteristics hold in a large number of industries, like the airline, financial services, restaurant or casino industries. To take an example from the financial services industry: A private banking advisor's performance is often random. A customer may use the performance stream to assess the account's likely long-run return and volatility, which may then influence his/her decision on whether to continue doing business with the firm. The firm observes the performance stream and can make different offers to different customers, potentially customizing according to the performance stream experienced by a customer and according to the customer's behaviors in response to those performances. In airlines, flight delays are random and firms typically record the delays experienced by each passenger. Whenever a customer uses Uber, the driver-rider marketplace firm, the quality of the driver they receive is random. Passengers rate the quality of their experience which Uber can then use for targeting purposes.

A primary contribution of this paper is a framework and methodology that allows the manager to forecast the likely marketing response for any customer with any specific experience and behavior history, which in turn can be used to decide which customers to target for marketing. An important insight identified from this methodology is that the optimal targeting decision may depend on a customer's expectations of quality because of its influence on marketing responsiveness. For instance, a customer with early experiences that are of atypically low levels will likely form a belief distribution with a low expectation of experience quality from future visits and, on the basis of this belief, may reduce or altogether stop interactions with the firm. Attractive direct marketing offers targeted at such a customer can increase the customer's expected utility for a future visit, incentivize the customer to interact further with the client, improve the belief distribution based on the new experience, increase the likelihood of future visits, and increase the future profits of the firm from that customer. This suggests that if the marketing offers are costly and the firm can make the offers to only a limited set of customers, then it should target those customers whose future profits will be most increased by the offer, possibly those customers whose belief distributions can be improved the most, which may be those customers who have had atypically low levels of experience quality. However, this intuition should not be taken to imply that the firm should simply direct offers to the customers with the lowest experience qualities. One has to balance the cost of an offer against the benefits, which will depend on marketing responsiveness and the extent to which the offer will influence the customer's future behaviors and the firm's profit from those behaviors. In this paper, I present a model to do exactly that and provide evidence that incorporating a customer's beliefs of experience quality into the targeting decision can significantly increase a firm's profit.

To address the direct marketing decision problem in this important class of situations where the manager can react to a customer's observed experience and behavior history, I employ a Bayesian, dynamic learning framework. A customer starts with a prior belief distribution on the mean value of the experience quality level and updates this belief distribution according to new experiences accumulated with each interaction with the firm. For a rational consumer, the decision to engage in repeat interactions with the firm, thereby forming more accurate beliefs, will depend on the utility from the interactions, the value of the increased belief accuracy, and the consequent utilities from potential future visits. The model presented in this paper allows the manager to determine the extent to which direct marketing influences these utilities, the customer's consequent interactions with the firm and firms' consequent profits.

Because prior beliefs, marketing responsiveness and utility function parameters can vary from customer to customer, it is important that the model and estimation methodology allow for across-customer heterogeneity. Incorporating learning into a dynamic discrete choice model with forward looking customers is difficult because the optimal choice is the solution to a complex Bellman equation, with a correspondingly difficult likelihood function. Including unobserved heterogeneity makes the problem even more difficult. An important contribution of this paper is that it develops a tractable solution to this class of problems by combining a forward simulation algorithm with Markov Chain Monte Carlo ideas.

I illustrate this paper's framework and methodology in the context of the casino industry. I focus on the gaming industry for a few reasons. First, it is a moderately large industry in the United States: Gaming revenues are now at an historic peak at $67 billion, with nearly 1,000 casinos operating in 39 states (Oxford Economics, 2014). Oxford Economics estimates the U.S. gaming industry contributes nearly $240 billion to the national economy (2014).[1] The pervasiveness of the industry is as significant as its size; nearly one third of Americans have gambled at a casino within the past twelve months.

[1] This includes both tribal and commercial casinos. In 2012 71.6 million gamblers spent $37.34 billion in gaming revenues at commercial casinos. This is more than they did on movies, craft beer, and outdoor equipment combined. (American Gaming Association, 2013).

Therefore, insights into this industry may be valuable. Second, there is limited research on the impacts of casino marketing. This is due primarily to the difficulty of obtaining sensitive casino data rather than the lack of importance of marketing in the gaming industry. Third, it is behaviorally interesting because these are real gamblers responding to uncertainty, as opposed to lab participants. Fourth, an advantage of studying casinos is that gambling outcomes are exogenous. Even though casinos can control the overall house advantage and its distribution, the outcome of any one trip for any individual is independent of that individual's history and can take any value from a diverse set. This can greatly help in model identification. Finally, many casinos base marketing offer values on gamblers' past *expected* losses, and not their actual outcomes. I discuss this later in more detail but the implication is that much of the offer endogeneity is removed.

I use data on real gamblers to understand how their outcomes influence the time until their return trip. Specifically, I answer the following: How do past outcomes influence gamblers' beliefs on the house advantage and how can marketers use this information in their one-to-one targeting decisions? To estimate these impacts, I specify a dynamic learning model that incorporates the belief uncertainty into the utility specification. In the traditional random utility framework, consumers know the attributes of their choices perfectly. Learning models extend the traditional framework by recognizing that consumers may have incomplete information and thus make choices based on perceived rather than actual attributes (Ching et al., 2013). In this model, gamblers learn about the casino's house advantage by gambling at slot machines. Gamblers use their beliefs to form future cost expectations, which influences their decision to return to the casino. The uncertainty in these beliefs can also influence the decision to return. By fully modeling the learning process (rather than simply conditioning on the last outcome), the model permits gamblers to use their entire trip history when forming future cost expectations. The reduced-form evidence supports a full structural model of the learning process.

Academic researchers have shown considerable interest in decision making under uncertainty for decades, roughly starting with Tversky and Kahneman (1974). Meyer (1981) found that temporal variability increases the cost of information gathering, which suggests that variability comes with a premium. More recent work studies uncertainty in customer satisfaction (Bolton, 1998), service quality (Rust et al., 1999; Boulding et al., 1993), and product attributes (Erdem and Keane, 1996). An important difference between some this paper and some of these previous papers is that the focus of this research is on how a customer's belief (and the uncertainty of this belief) can influence a firm's targeted marketing actions. The intent is similar to that of Narayanan and Manchanda (2009), who also study learning behavior and its impact on targeting. However, a substantial econometric contribution of this research is that I allow gamblers to be forward looking. Gamblers can make tradeoffs between today's knowns and tomorrow's potential upsides to decide whether to continue engaging with the firm.

The dataset used in model estimation comes from a large casino in the United States. I observe the complete trip histories and marketing offers for a random sample of gamblers. The empirical strategy takes two parts. First, I show descriptive and reduced-form evidence that motivate the need for a structural learning model. Gamblers who incur a single loss return to the casino about ten days later than gamblers who incur a single win, but the return time increases as more losses are incurred: gamblers with four past losses return about thirty days later than a gambler with four past wins. These findings suggest that gamblers incorporate outcomes from multiple past trips into the decision process. Reduced-form evidence also shows gamblers' return times are significantly influenced by their beliefs about the house advantage and the uncertainty surrounding their beliefs.

Next, I estimate a structural model of the return time using a dynamic discrete choice framework. As is widely recognized, structural methods allow for counterfactual predictions about how changes in marketing policies will affect consumer behavior (Reiss, 2011). Our structural model incorporates the dynamic forward-looking behavior of individuals. One obstacle to estimating such structural models has been the contingent computational burden, which is mainly due to two reasons. First, the likelihood is based on the explicit solution of a dynamic programming (DP) model. This requires us to obtain the fixed point of a Bellman operator for each possible point in the state space. Second, the number of points in the state space increases exponentially with the dimensionality of the state space, a phenomenon referred to as the "curse of dimensionality". Imai et al. (2009) introduce a full-solution Bayesian approach to estimation of structural parameters. An important innovation in their algorithm is that they only need to conduct a single iteration of the Bellman operation during each estimation step (i.e., each MCMC draw). While conventional methods estimate the model only after solving the DP problem, their approach simultaneously solves the DP problem during parameter estimation. Because of this, the computational burden of their method is similar to that of non-Bayesian approaches but still intractable for dynamic learning models. In this paper I use forward simulation (see Bajari et al. (2007) and Hotz et al. (1994)). This significantly reduces the computation time and makes Bayesian estimation of a complex learning model feasible.

Learning models were first applied to marketing by Roberts and Urban (1988) and Eckstein et al. (1988). The initial models were relatively simple due to limitations on computer processing speeds and estimation algorithms available at the time. Erdem and Keane (1996) represents a significant methodological advance because it expanded the the class of learning models that became feasible to estimate. They used the method of Keane and Wolpin (1994) to obtain a fast and accurate approximate solution to the dynamic optimization problem and used simulation methods to approximate the likelihood function.

Recently, a few research papers applied the modified Bayesian MCMC algorithm first proposed by Imai et al. (2009) to estimate learning models with forward-looking consumers (see Roos et al. (2013)). Osborne (2011) is the first paper to allow for both learning and switching costs as sources of state dependence in a forward looking learning model. He also incorporated unobserved heterogeneity, however his paper assumed a "one-shot" learning model where only one purchase occurrence is all that is needed to learn everything about the product. While this is defensible in the product category he analyzed (laundry detergent), the purchase to purchase variability in many settings, including gambling detracts from the basic assumption underlying this method. The model considered in our paper allows for learning to happen over multiple exposures and also allows for individual parameters to be estimated using Bayesian methods with flexible mixture distributions on the heterogeneous parameters.

Our empirical results suggest that gamblers' prior beliefs translate to an overestimation of the house advantage by a factor of about four. Further, our counterfactual analyses suggest that this overestimation may be costing the casino substantial amounts of revenue. When gamblers overestimate the house advantage, they overestimate projected future expenditures, which reduces the probability of revisiting the casino within any given time period. The counterfactuals also illustrate the value of incorporating the beliefs and outcomes into the targeted marketing decision. I show that naive targeting strategies based on simple outcome heuristics are not sufficient and that more sophisticated targeting strategies can lead to higher revenue. The final simulation shows that marketing strategies which vary offer amounts based on gamblers' beliefs in the house advantage improve profitability by close to 20%.

1.1.1 The Casino Industry

The gaming industry is a critical component of the U.S. economy. Casino gaming revenues are now at an historic peak at $67 billion. Oxford Economics (2014) estimates the U.S. gaming industry contributes nearly $240 billion to the national economy. With nearly 1,000 casinos operating in 39 states, Americans spend more money enjoying casino entertainment than they spend on spectator sports like football, baseball, basketball and soccer combined. Despite its substantial contribution to the economy, there are only a handful of papers that study effects of casino marketing (Nair et al., 2013; Park and Manchanda, 2015; Narayanan and Manchanda, 2012).

The casino industry has long understood the importance of effective customer relationship management (Compton, 1999). In today's gaming environment, a sophisticated tracking system is essential to remain competitive, especially in saturated markets like those of Las Vegas and Atlantic City (Kilby and Fox, 1998). Casino marketing offers typically include a combination of free room nights (if the casino has a hotel) and slot promotional credits. Offers can also include additional complimentaries (or "comps") for virtually any other amenities available at the casino, such as shows, spa treatments, or dining.

To determine the optimal level of comps to offer their players, managers need to estimate gaming revenue at the player level. Casinos use player rating systems to track and record individual plays and player information. The marketing department uses these data to segment customers into tiers for targeting marketing offers.

Casinos record player information by having them enroll in the casino's loyalty program. Most casinos have enrollment centers on the casino floor. To incentivize gamblers to have their play tracked, casinos offer rewards programs. Rewards programs are different across casinos, but one feature they share is that the magnitude of the reward is a function of play volume and possibly on-property purchases. A rewards system increases the likelihood that the casino has the complete play history recorded because without rewards players are unlikely to allow the casino to track their play. While no-card play (that is, gambling without a loyalty card) still contributes a substantial portion to a casino's revenues, the amount of no-card play per person is often insignificant and unlikely to be of interest to the casino. Frequent gamblers typically understand that it is in their best interest to have all of their play tracked to increase rewards earning.

Slot Machines I limit our empirical analysis to gamblers who play only slot machines, for three reasons. First, electronic gaming machines are the most popular game among casino visitors, as more than half (51 percent) choose slot machines or video poker as their favorite game (American Gaming Association, 2013). Second, tracking table games activity at the individual player level is still a very manual process and often inaccurate. On slot machines, all play is recorded electronically through each gambler's loyalty card and, because of this, revenue is exact down to the penny. Third, the skill of a tablegame player can to some degree dictate the outcome. For example, a skilled blackjack player can reduce the house advantage to nearly zero (or negative if they are counting cards) but an unskilled player can lose far more than expected in the long run. On slot machines it does not matter who is pulling the lever (or nowadays more often pressing the "spin" button); the outcome is completely random and in the long run the hold percentage should converge to the house advantage regardless of the individual gambler.

Before proceeding further, I introduce a few terms and concepts which are used by industry practitioners and by us in the exposition of our learning model. *Handle* measures the total amount of money wagered on the machine. This measure of volume allows management to monitor the overall popularity of games. The *hold percentage* is the percentage of handle that the machine keeps in any particular play event. For example, if a player plays puts in \$5 and gets out \$6, then the hold is -20%, but if the player loses all of the \$5 then the hold is 100%. The hold percentage can vary quite a lot across play events, and is governed by the randomness programmed into the the particular machine. Different machines can have different hold distributions programmed into them. The hold percentage for a consumer on a certain trip is the percentage of that trip's play volume lost to the casino, and is an aggregation of the hold percentages of the machines that he/she played on. The *house advantage* is the expected hold percentage of the slot machine and depends on the payouts and odds specific to that machine. The advantage can also be seen, based on the weak law of large numbers, as the hold percentage aggregated over a large number of plays. Slot machine advantages range from as low as .5% to as much as 25%. Many gaming jurisdictions have established minimum levels at which slot machines must pay back in order to prevent casino operators from placing too great a disadvantage on players (Kilby and Fox, 1998). The Las Vegas Strip house advantage is around 7%, in Reno it is about 5%.[2] The important point to remember is that the hold percentage reflects the *actual* empirically realized amount of money kept by the casino, while the house advantage is what the casino *expects* to keep.

Theoretical and Actual Outcomes Casino operators track both actual and theoretical player losses. The theoretical loss (also called "theo") is the amount of money the player was expected to lose. It is based on the following formula:

$$\begin{aligned} \text{Theoretical Loss} &= \text{Avg. Bet} \times \text{Decisions per Hour} \times \text{Hours} \times \text{House Advantage} \\ &= \text{Handle} \times \text{House Advantage} \end{aligned}$$

As mentioned earlier, for targeting purposes casinos typically ignore the actual outcomes and instead value their players on theoretical losses alone. The primary reason for doing so is to control for the randomness of outcomes. This creates a significant advantage for an analyst studying the impact of outcomes because marketing offer values are not endogenous, in the sense that the offers are not directly tied to the actual outcomes. For example, if one gambler loses \$500 and the another wins \$4,000 but they were both *expected* to lose \$300 they will receive the same offer.

1.2 Data

The dataset used to estimate the structural model comes from a large casino in the United States. The dataset includes the complete trip histories from over 28,000 randomly selected slot gamblers with around 110,000 trips occurring between February 2006 and May 2015. I observe basic demographic information such as gender, distance to the casino,

[2] http://gaming.unlv.edu/reports/nv_slot_hold.pdf

age, and current loyalty card level (either "Silver" or "Gold").[3] A "trip" is defined as a distinct period of time where gambling activity is observed. For instance, a new trip is initiated when either 1) a player inserts their loyalty card into a slot machine for the first time or 2) a significant lapse in play occurs. The lapse used to demarcate a new trip is set by management in a way that makes it very likely that each trip record captures a distinct return to the casino rather than simply a suspension in play within a single trip. Typically, the cutoff is three days, meaning that if no activity is observed for three days, the trip is considered to have ended and any further play initiates a new trip record.[4]

Each trip record includes detailed information about the gambling activity from that trip. The variables of interest are the start and end dates of the trip, actual and theoretical loss values, time played, average bet, promotional credits redeemed, comps received, and whether they stayed at the hotel or not. I also observe all marketing activity for these gamblers. Over the observed period, the gamblers redeemed over 2,500 separate offers. For each offer, I observe the period for which the offer is active (typically about 2.5 months), the date when the offer was sent to the gamblers (the "drop" date), the total number of promotional credits in the offer, and the comp room type. A slot promotional credit is essentially free slot play where any winnings from the promotional credit can be kept. The promotional value itself cannot be converted to cash.

1.3 Descriptive Analysis and Reduced-Form Evidence

Before discussing the structural estimation procedure, I will first describe the data and show reduced-form evidence that gambling outcomes impact the timing of the return trip. I exclude players whose play is high enough to warrant a casino host. This ensures that the only marketing communication between the casino and the gambler is done through direct marketing offers. I only keep gamblers who play slot machines exclusively, for

[3]Since I cannot reveal the actual labels used by the anonymous casino, I will refer to the upper loyalty tier as the "Gold" level.

[4]After the three-day lapse, the trip is recorded as "ending" on the last day of play, not three days later.

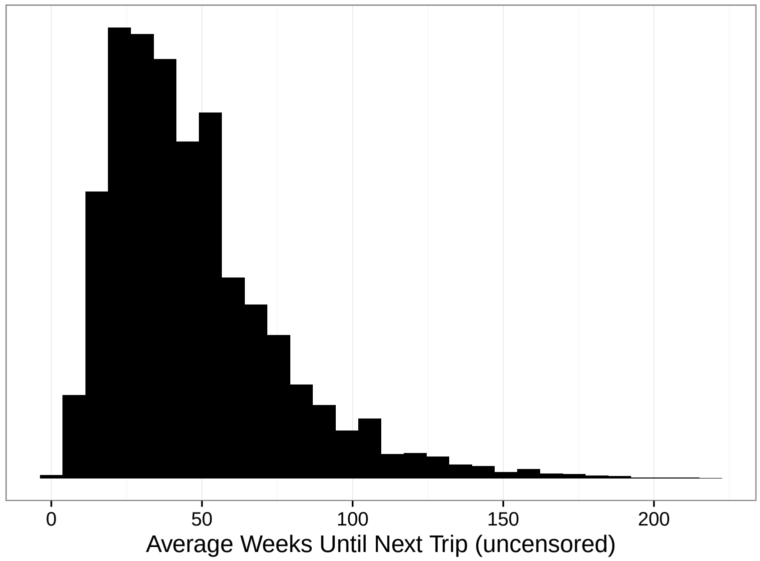
reasons discussed previously, which represents 37% of the low-end player base. I remove gamblers whose first trip to the casino occurred before the first available marketing offer data is observed so that I have the complete marketing and trip history for each gambler. Table 1 summarizes the cleaned data. For estimation, I only use gamblers with at least three trips to the casino. This is done to ensure that each gambler observes a sufficient amount of variance in the experienced hold percentages. In general, their statistics are quite similar to the aggregate level statistics.

Table 1: Descriptive Statistics

	All Data	Gamblers >= 3 Trips
Gamblers	28,362	13,964
Trips	113,752	94,139
Trip Length	2.5 days	2.8 days
Slot Average Bet	$1.95	$1.69
Spins/Minute	7.7	7.8
Hours Played/Trip	5.3	5.9
Hold % Experienced	12%	10%
Expected Loss/Trip	$387	$416

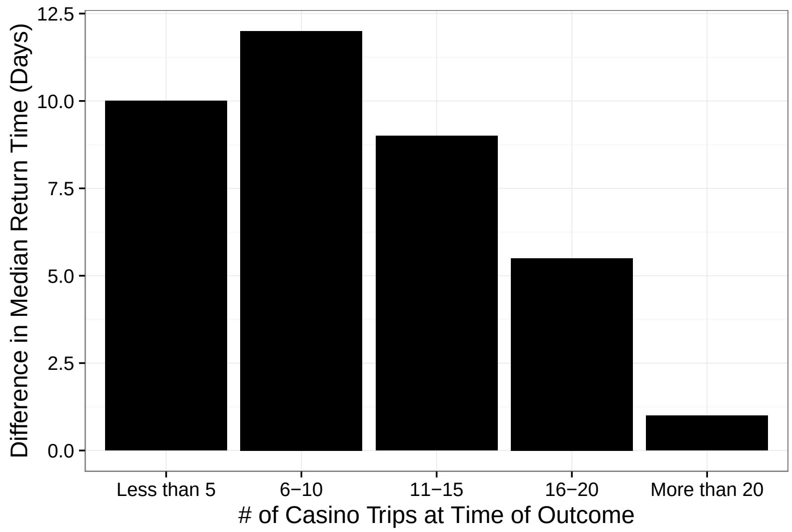
Figure 1 shows the distribution of average uncensored intervisit times across gamblers with at least three trips at the casino. The median return time is about ten months.

Figure 1: Intervisit Timing Distribution



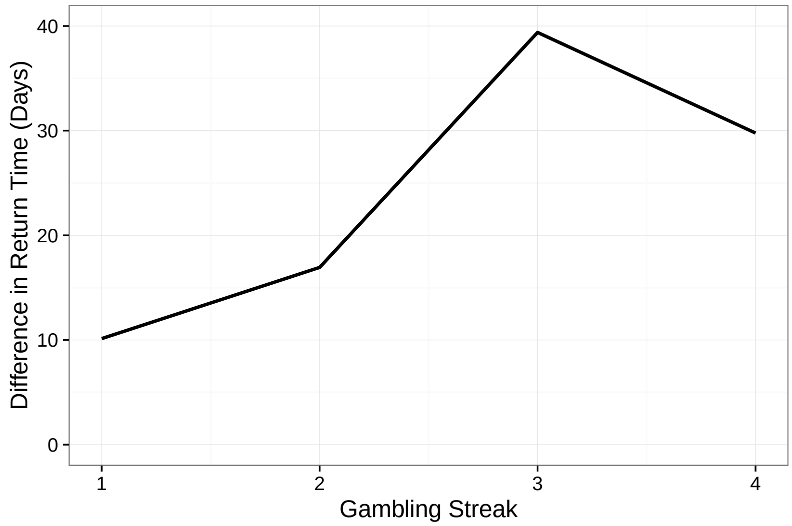
If gamblers learn from experience, their sensitivity to a single trip's gambling outcomes should decline over time. Experienced gamblers have more certainty about expected outcomes and therefore should be less likely to be swayed by their most recent outcome. Inexperienced gamblers (those with only a few trips) project expected outcomes only using a few signals, which can vary greatly between players and cause biases depending on the sequence realized. Over time gambler beliefs will converge to the truth and a single outcome will not have as much of an impact on the return decision. I see evidence for this in the data. Figure 2 shows the difference in median return times after a gambler lost compared with after a gambler won. The differences are grouped by experience, represented by the number of trips to the casino when the win or loss was realized. For example, when gamblers have less than five casino trips, they tend to return about ten days later when they lose versus when they win. With more trips (and more experience) the difference in return times diminishes and gamblers become less impacted by the most recent trip's win or loss. To handle selection bias (that is, gamblers who eventually have many trips at the casino are inherently different from those with fewer trips), I only include players who eventually have between 15 and 30 total trips at the casino.

Figure 2: Median Return Lag After a Player Loss (Relative to a Player Win)



Additional evidence for learning across trips is given by Figure 3 in which I show the difference in return times between winning and losing streaks, where the streak occurs over the past one through four trips. Ignoring any streaks (where the gambling streak equals one), players that lose tend to come back 10 days later than those that win. Moreover, as players lose multiple times in a row, they delay the return trip by even greater lengths. For example, a gambler who lost three trips in a row will return about forty days later compared with a gambler who won three trips in a row.

Figure 3: Return Lag by Player Losing Streak (Relative to Player Winning Streak)



Next I provide reduced-form evidence that gamblers' return times to the casino are influenced by their beliefs in the house advantage. I estimate return times using a Weibull hazard model and include the posterior mean of the house advantage and its posterior variance as covariates – the posteriors are generated from a Bayesian learning process (the updating process is discussed in detail later). I also include several demographic and last-trip variables: age, sex, card level, distance to the casino, whether they stayed at the hotel on the last trip, whether they redeemed a promotion on the last trip, the log of the total comps received, trip length, and last trip theoretical loss.

In a Bayesian learning process, the prior mean and prior variance dictate the evolution of the posterior mean and variance. Because of this, in a reduced-form hazard model a prior mean and variance needs to be selected in order to generate the posterior beliefs on the house advantage (the mean and its uncertainty). In the structural model, these priors are estimated, but for reduced-form evidence I estimate 150 hazard models over a grid of 15 prior mean and 10 prior variance starting points. Figure 4 shows examples of the truncated normal shapes to illustrate the variety of prior settings that are considered for the house advantage beliefs. The idea is that by estimating many hazard models over this grid I can determine if the reduced-form coefficient estimates are sensitive to the learning process priors. The specific gridpoints are available in the Appendix.

Figure 4: Truncated Normal Distribution Examples Using Gridpoints

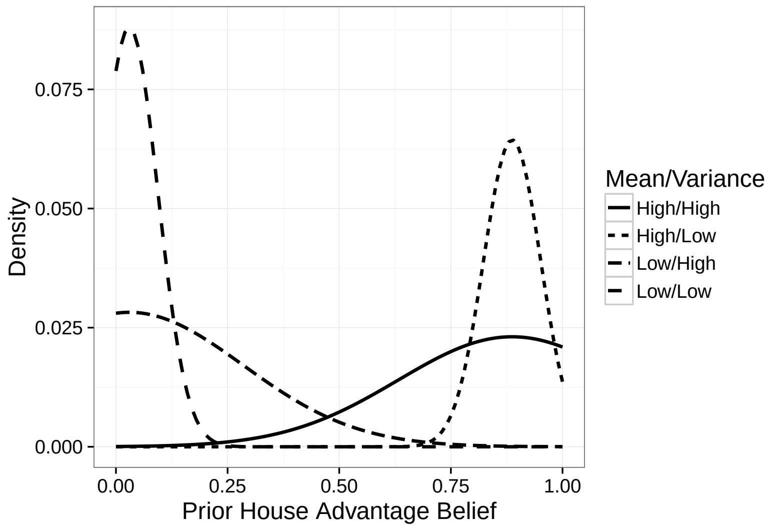


Figure 5 plots the coefficient on the posterior mean across all 150 gridpoints. The latticed plane is positioned where the z-axis equals zero; any points above this plane are positive and below are negative and points that are filled in are significant at the .05 level. Except for very low prior variance values (where convergence of weibull is not achieved) the coefficients on the posterior mean tend to be positive and significant, which means that the return time increases with the posterior mean of the house advantage. In other words, as the belief in the house advantage increases gamblers take longer to return. Figure 6 shows a similar plot but for the coefficient on the posterior variance. At very low prior variance settings the coefficient cannot be estimated. The coefficients that can be estimated are significant, suggesting that uncertainty in the gamblers' beliefs on the house advantage influence the return decision.

The reduced-form evidence suggests that 1) learning should be incorporated into a model of the return times, and 2) posterior beliefs in the house advantage (the mean and its uncertainty) influence the return time. The drawback of a reduced-form approach is that it does not account for any forward-looking behavior of the gamblers.

Figure 5: Weibull Posterior Mean Coefficients

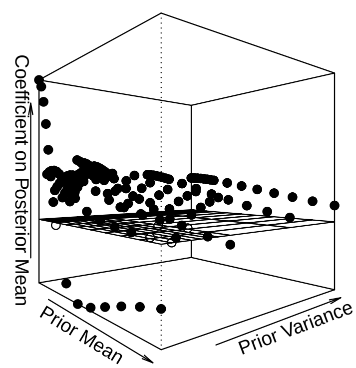
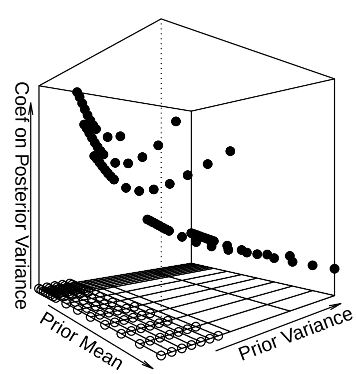


Figure 6: Weibull Posterior Variance Coefficients



1.4 A Model of Gambler Learning

In this section, I propose a structural model of the casino return decision process. I first outline the dynamic optimization problem somewhat generally and then introduce the learning component and specific utility function.

1.4.1 Gambler Dynamic Optimization Problem

I model casino return times in the framework of a dynamic discrete choice model, which can be interpreted as a generalization of a structural hazard model Van den Berg (2001). I estimate an infinite horizon model of a forward looking agent. In each decision period, the gambler decides to return to the casino or not by comparing his/her current and discounted future utilities of each action.

Let θ be the J-dimensional parameter vector. Let S be the finite set of state space points and s be an element of S. Let A be the finite set of all possible actions and a be an element of A. Let $u(s, a, \varepsilon_a, \theta)$ be the current period utility of choosing action a , given state s and ϵ , is a vector whose ath element is a random choice to the current returns of choice a. The transition probability of next period state s', given current state s and action a is $f(s'|s, a, \theta)$. Given a discount rate β, The time invariant value function can be defined to be the maximum of the discounted sum of expected utilities:

$$V(s_t, \varepsilon_t, \theta) \equiv \max_{\{a_t, a_{t+1}, \ldots\}} \mathbb{E}\left[\sum_{\tau=t}^{\infty} \beta^{\tau} u(s_\tau, a_\tau, \epsilon_{a_\tau}, \theta) | s_t, \varepsilon_t\right]$$

This value function is known to be the unique solution to the Bellman equation:

$$V(s, \varepsilon, \theta) = \max_{a \in A} \{\mathbb{E}[u(s, a, \varepsilon_a, \theta)] + \beta \mathbb{E}_{s', \varepsilon'}[V(s', \varepsilon', \theta) | s, a]\}$$

The first expectation is included because even when making the decision to return to the casino the utility is not known until after the trip has been realized. The second expectation is taken with respect to the next period shock ε' and the next period state s'.

If I define $EV(s, a, \varepsilon_a, \theta)$ to be the expected value of choosing action a then

$$EV(s, a, \varepsilon_a, \theta) = \mathbb{E}\left[u(s, a, \varepsilon_a, \theta)\right] + \beta \mathbb{E}_{s', \varepsilon'}\left[V(s', \varepsilon', \theta) | s, a\right]$$

and the value function can be written as

$$V(s, \varepsilon, \theta) = \max_{a \in A} EV(s, a, \varepsilon_a, \theta)$$

The dataset for estimation includes variables which correspond to state vector s and choice a but the choice shock ε is not observed. I observe data for $i = 1, \ldots, N$ gamblers, and each gambler i has T_i observations. The observed data for individual i is denoted $y_i^d \equiv \left\{a_{i,t}^d, s_{i,t}^d\right\}_{t=1}^{T_i}$ and $Y^d \equiv \left\{y_i^d\right\}_{i=1}^{N}$ with superscript d to represent that this is observable data. Furthermore,

$$a_{i,t}^d = \arg\max_{a \in A} EV\left(s_{i,t}^d, a, \varepsilon_a, \theta\right)$$

Let $\pi(\cdot)$ be the prior distribution of θ and let $L\left(Y^d | \theta\right)$ be the likelihood of the model, given the parameter θ and the value function $V(\cdot, \cdot, \theta)$, which is the solution of the dynamic programming problem. Then I have a posterior distribution function of θ:

$$P\left(\theta | Y^d\right) \propto L\left(Y^d | \theta\right) \pi(\theta)$$

Let $\varepsilon_i \equiv \{\varepsilon_{i,t}\}_{t=1}^{T_i}$ and $\varepsilon \equiv \{\varepsilon_i\}_{i=1}^{N}$. The expressions above are conditional on ε. Because ε is not observed to the analyst, the unconditional likelihood needs to be used, obtained by integrating over it. That is, if I define $L\left(Y^d | \varepsilon, \theta\right)$ to be the likelihood conditional on (ε, θ), then

$$L\left(Y^{d}|\theta\right)=\int L\left(Y^{d}|\varepsilon,\theta\right)dF_{\varepsilon}\left(\varepsilon|\theta\right)$$

The value function enters into the likelihood through the choice probability. The period-specific utility is written as follows:

$$u\left(s,a,\varepsilon_{a},\theta\right)=\widehat{u}\left(s,a,\theta\right)+\varepsilon_{a}$$

, where $\widehat{u}\left(s,a,\theta\right)$ is the deterministic component of the per-period utility. Furthermore,

$$\widehat{EV}\left(s,a,\theta\right)=\mathbb{E}\left[\widehat{u}\left(s,a,\theta\right)\right]+\beta\sum_{s'}\mathbb{E}_{\varepsilon'}\left[V\left(s',\varepsilon',\theta\right)\right]f\left(s'|s,a,\theta\right)$$

This leads to:

$$\Pr\left[a_{i,t}^{d}|s_{i,t}^{d},V,\theta\right]=\Pr\left[\varepsilon_{a}-\varepsilon_{a_{i,t}^{d}}\leq\widehat{EV}\left(s,a_{i,t}^{d},\theta\right)-\widehat{EV}\left(s,a,\theta\right);\forall a\neq a_{i,t}^{d}|s_{i,t}^{d},V,\theta\right]$$

I assume that each ε_a is independently drawn from the same extreme value distribution. In addition, I introduce a hierarchical structure so that each V and θ are specific to gambler i. In the empirical application since A contains two actions, either return to the casino $(a=1)$ or not $(a=0)$, the conditional choice probabilities take the following form:

$$\Pr\left[a_{i,t}^{d}=1|s_{i,t}^{d},V_{i},\theta_{i}\right]=\frac{1}{1+\exp\left(-\left[\widehat{EV}\left(s_{i,t}^{d},1,\theta_{i}\right)-\widehat{EV}\left(s_{i,t}^{d},0,\theta_{i}\right)\right]\right)}$$

1.4.2 Introducing a Hierarchical Structure

To allow for parameter estimates to vary by individual characteristics, I introduce a hierarchical structure. Understanding individual differences is crucial in strategic CRM applications when developing targeted marketing strategies (Rossi et al., 2005). The hierarchical parameters are specified as a function of an individual's observable characteristics.

I have nz observable characteristics on each individual. If I let Z denote a matrix with N rows and nz columns and similarly Θ be a matrix of N rows and J columns, where the ith row of Θ is the parameter estimates for individual i then I have:

$$\Theta = Z\Delta + U$$

Where Δ is a $nz \times J$ matrix of coefficients on the observables and U is a vector of residuals. This is simply a multivariate regression of Θ on Z. In each row of U,

$$u_i \sim N\left(0, \Sigma_\theta\right)$$

The priors are specified as follows:

$$\begin{aligned} \operatorname{vec}\left(\Delta|\Sigma_\theta\right) &\sim N\left(\operatorname{vec}\left(\bar{\Delta}\right), \Sigma_\theta \otimes A^{-1}\right) \\ \Sigma_\theta &\sim IW\left(\nu, \Sigma\right) \end{aligned}$$

Hierarchical models for panel data structures are ideally suited for MCMC methods. A Gibbs-style Markov chain can be constructed by considering the two sets of conditionals:

$$\theta_i|\tau, y_i^d$$
$$\tau|\left\{\theta_i\right\}$$

The first line exploits the fact that the θ_i are independent, conditional on the first stage priors $\tau = \{\Delta, \Sigma_\theta\}$. The second line exploits the fact that $\{\theta_i\}$are sufficient for τ. That is, once the individual level parameters are drawn they serve as "data" to the inferences on the priors. Due to the non-linearity of the model proposed, there is no way to conveniently sample from the conditional posterior (i.e., using a Gibbs sampler).

For this reason, I employ a Metropolis algorithm to draw θ_i For each gambler i, I draw candidate random effects parameters θ_i^n by perturbing the current draw θ_i^o: $\theta_i^n = \theta_i^o + \varepsilon$, where $\varepsilon \sim \mathcal{N}\left(0, s^2\Sigma\right)$. I then compare the likelihood of the new parameters with the old parameters and accept the new parameters with probability α:

$$
\begin{aligned}
\alpha\left(\theta_i^n, \theta_i^o\right) &= \min\left\{1, \frac{L_i\left(Y_i^d|\theta_i^n\right) q\left(\theta_i^n, \theta_i^o\right)}{L_i\left(Y_i^d|\theta_i^o\right) q\left(\theta_i^o, \theta_i^n\right)}\right\} \\
&= \min\left\{1, \frac{L_i\left(Y_i^d|\theta_i^n\right)}{L_i\left(Y_i^d|\theta_i^o\right)}\right\}
\end{aligned}
$$

The second line is a result of the symmetry of the transition density $q\left(\cdot, \cdot\right)$.

1.4.3 Learning About the House Advantage

In this section, I introduce the learning process. As gamblers play slot machines, they receive signals about that casino's house advantage. Before receiving any information, they have a truncated normal prior belief on the house advantage:

$$
A_i \sim \mathcal{TN}\left(A_{0i}, \sigma_{0i}^2, 0, 1\right)
$$

In other words, before a gambler's first trip to the casino, they expect to lose a certain percentage of every dollar cycled through the slot machine. The house advantage is bounded from below at zero because it is irrational for a gambler to expect to win money from a slot machine in the long run. It is also bounded from above at one because in the long run it is impossible for a machine to pay out more money than is put into it. Again, in the short term the *hold percentage* can fall outside of these bounds, but the gambler's beliefs on the *house advantage* cannot reasonably be outside of this range.

The player's experience at the casino does not fully reveal the house advantage because of the inherent variability of gambling outcomes. As previously discussed, there is quite often a difference between the *hold percentage* and the *house advantage* for gambler i

on occasion t. I denote the hold percentage as H_{it}, which can be interpreted as the "experienced" house advantage, and the house advantage as A_i. The hold percentage is thus a noisy signal of the house advantage:

$$H_{it} = A_{it} + \eta_{it}, \text{ where } \eta_{it} \sim \mathcal{N}\left(0, \sigma_{\eta}^{2}\right)$$

The gamblers update their posterior mean and variance of the house advantages using a Bayesian updating process. The updating formulas are given below:

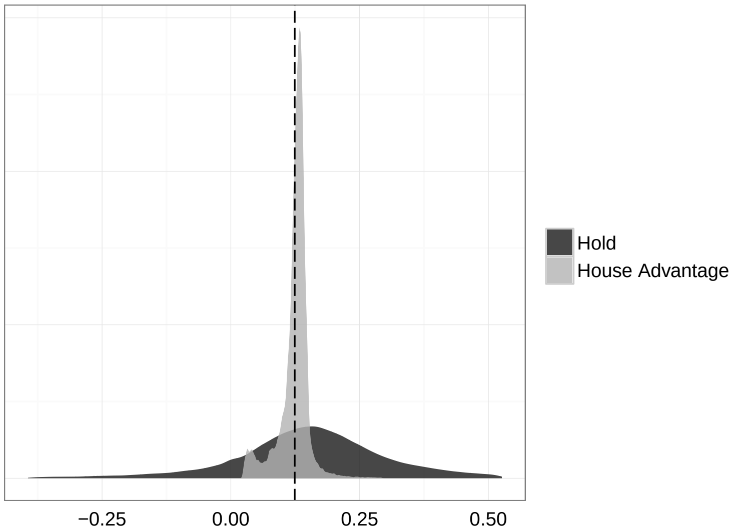
$$A_{it} = \frac{\sigma_{0i}^{2}}{N_i(t) \cdot \sigma_{0i}^{2} + \sigma_{\eta_i}^{2}} \sum_{s=1}^{t} H_{is} d_{is} + \frac{\sigma_{\eta_i}^{2}}{N_i(t) \cdot \sigma_{0i}^{2} + \sigma_{\eta_i}^{2}} A_{0i}$$

$$\sigma_{it}^{2} = \frac{1}{1/\sigma_{0i}^{2} + N(t) \cdot 1/\sigma_{\eta_i}^{2}}$$

where $N_i(t)$ is the number of gambling experiences realized up through time t and d_{it} is an indicator for whether the player gambled at time t. The Appendix contains a proof showing that if the prior is truncated normal and the signal is an unbounded normal then the corresponding posterior is also a truncated normal.

Figure 7 plots the distribution of the hold percentage and house advantages across all trips in the dataset. The hold percentage distribution is what the gamblers experience and the house advantage is what gamblers are attempting to learn about. The dashed vertical line is the mean house advantage: with a sufficient number of exposures, the gamblers will learn this value with certainty if slot machines are selected at random. Note that each gambler's experienced house advantage is observed by the analyst, so even if gamblers do not select slot machines at random (for instance, they only play one machine that happens to have a very low house advantage) the analyst can still determine if their estimated posterior beliefs are above or below the true house advantage. However, because the distribution of the house advantages is so tightly centered, I make the defensible simplifying assumption that the machines are selected at random and only the mean house advantage matters.

Figure 7: Gamblers Learn the House Advantage from the Hold Percentages Experienced



1.4.4 Cost of Gambling

When gamblers consider a return trip to the casino they need to form projections on the cost of gambling. This influences the expected future utilities. Under perfect knowledge expected cost is the same as the theoretical loss:

$$\text{Gambling Cost} = \text{Avg. Bet} \times \text{Decisions per Hour} \times \text{Hours} \times \text{House Advantage}$$

However, since gamblers have imperfect knowledge on the house advantages there is uncertainty in projections of their gambling costs. This uncertainty depends on their current beliefs at time t:

$$\text{Gambling Cost} \sim \mathcal{N}\left(\text{BDH} \cdot A_{it}, \text{BDH}^2 \cdot \sigma_{it}^2\right)$$

"BDH" represents the product of average bet, decisions per hour, and hours played. These three variables are completely within the gambler's control (i.e., there is no uncertainty) and represent the gambler's play style.

It is important to note that the gambler's projected average bet, game speed, and time may be a function of their current beliefs on the house advantage. For example, if faced with a relatively high house advantage players may decide to decrease their average bet to reduce projected gambling costs (everything else held constant). Similarly, higher uncertainty in the cost may lead to play that is more likely to result in a lower cost. Furthermore, gamblers may also adjust their play style based on currently available marketing offers. For instance, a gambler returning to the casino on a free room offer may play more aggressively than usual since the comped room frees up money that could be used for gambling. To account for this, the BDH value can vary during the forward simulation (discussed in more detail later).

1.4.5 Utility Specification

In this section I introduce the utility function. The utility associated with returning to the casino is given by the following expression:

$$\begin{aligned}
u\left(a=1, s, \varepsilon_1, \theta\right)_{it} &= \theta_{1i}\left(\mathrm{BDH}_{it} \cdot H_{it}\right) \\
&\quad +\theta_{1i} r_i \left(\mathrm{BDH}_{it} \cdot H_{it}\right)^2 \\
&\quad +\theta_{2i}\text{Offer Gaming Value}_{it} \\
&\quad +\theta_{3i}\text{Offer Room Value}_{it} \\
&\quad +\Omega f\left(w_{it}\right) + \Gamma\text{Month}_{it} + \theta_{0i} + \varepsilon_{1it} \\
u\left(a=0, s, \varepsilon_0, \theta\right)_{it} &= \varepsilon_{0it}
\end{aligned}$$

Where u_{it} is the utility for gambler i at time t. BDH is the product of average bet, pulls per hour, and hours played.[5] BDH multiplied by the hold H it captures the gambling expense realized from that trip. Importantly, this expense is not known at the time of the decision and only realized after experiencing the outcome. θ_1 represents the utility

[5]Note BDH is the same as the handle, but in order to prevent confusion between handle and hold I call it BDH

weight gamblers attach to this cost, r is the risk coefficient, θ_2 is the utility weight of the offer's gaming value, and θ_3 is the utility weight of the offer's room value. Ω is the vector of utility weights associated with a function of the time since the last trip (w), which I specify as polynomials: $\Omega f(w) = \omega_1 w + \omega_2 w^2 + \omega_3 w^3 + \omega_4 w^4 + \omega_5 w^5$. Γ is the vector of utility weights associated with the month the decision to return was made, in order to capture impacts from seasonality: $\Gamma\text{Month} = \gamma_1 \mathbb{I}[\text{Month} = 1] + \ldots + \gamma_{11} \mathbb{I}[\text{Month} = 11]$. θ_0 is an intercept. ε is the random component associated with this choice, which is known to the gambler but not observed by the analyst. Ω and Γ are common across individuals, while θ_0, θ_1, θ_2, and r are specific to the individual.

Given the utility specification and the learning process, expected utility is given by the following:

$$
\begin{aligned}
\mathbb{E}_{A_{it}}\left[u\left(a=1, s, \varepsilon_1, \theta\right)\right] &= \theta_{1i}\left(\text{BDH}_{it} \cdot A_{it}\right) + \theta_{1i} r_i \left(\text{BDH}_{it} \cdot A_{it}\right)^2 + \theta_{1i} r_i \left(\text{BDH}_{it}^2 \cdot \left(\sigma_{it}^2 + \sigma_{\eta i}^2\right)\right) \\
&\quad + \theta_{2i} \text{Offer Gaming Value}_{it} \\
&\quad + \theta_{3i} \text{Offer Room Value}_{it} \\
&\quad + \Omega f\left(w_{it}\right) + \Gamma \text{Month}_{it} + \theta_{0i} + \varepsilon_{1it} \\
u\left(a=0, s, \varepsilon_0, \theta\right) &= \varepsilon_{0it}
\end{aligned}
$$

Under this specification, utility is linear in the cost of gambling. As in Erdem and Keane (1996), the formulation is such that given a strictly negative θ_1, utility is concave in A for $r > 0$, linear in A for $r = 0$, and convex for $r < 0$. Thus if there is uncertainty about the house advantage, the consumer is risk averse, risk neutral or risk seeking as $r > 0$, $r = 0$, or $r < 0$, respectively. As noted earlier, the uncertainty is in the beliefs on the house advantage, even in the "current" decision period. Furthermore, while the offer values are known in the current period they are not known in future periods, so gamblers form expectations over these values as well. In the simulation I draw values from the empirical joint distribution of room and gaming offer values.

1.5 Model Estimation

1.5.1 The Estimation Procedure

The structural parameters of interest are $\{\theta_{0i}, \theta_{1i}, \theta_{2i}, \theta_{3i}, r_i, \Omega, \Gamma\}$ and the priors on each individual's learning process $\{A_{0i}, \sigma_{0i}^2\}$. The proposed estimation procedure uses the advantages of Bayesian estimation (versus classical estimation methods) while remaining computationally feasible. The biggest challenge presented when estimating structural learning models is that the state space is incredibly large. When discounting future expectations, a forward looking gambler needs to consider the impacts of all potential outcomes and the associated implications on the learning process itself. For example, the specific hold percentage a gambler expects to experience on a return trip will influence how their posterior beliefs update, which in turn influences later return decisions. Clearly, evaluating every single potential learning path is daunting and because of this a full-solution Bayesian approach is not feasible, such as the method proposed by Imai et al. (2009).

Erdem and Keane (1996) use backwards induction to solve their learning model. However, the entire backwards induction needs to be re-solved at every parameter estimate. This is not feasible for Bayesian methods, which typically rely on tens of thousands of MCMC draws to converge onto the posteriors. The impracticality of their method is not limited simply to the desire to use Bayesian rather than classical methods: the complexity of the proposed utility function and the hierarchical structure also render their approach as unfeasible.

Rather than attempt to visit every single learning path I forward simulate over many potential paths and discount the simulated values. More likely paths will be simulated more often and averaging over many simulated paths provides a consistent estimate of the discounted future returns. The advantage of this approach is that if the utility function is linear in parameters I only have to simulate the paths once for each considered starting state since the current parameter estimates do not affect the simulated values (see Hotz et al. (1994) for a discussion of this method). The discounted terms can be separated from the parameters such that the parameters simply scale the discounted values during estimation.

One challenge is that in the utility function specified some variables enter non-linearly, namely the prior mean and prior variance of the beliefs in the house advantage. Note that the variability in the hold percentage ($\sigma^2_{\eta_i}$) is observed by both the analyst and the gambler, so there is no need to estimate this. To handle the non-linearity of the learning priors, I forward simulate over a grid of prior mean and prior variance values and during likelihood evaluation use bi-linear interpolation to fill in areas near the simulated prior mean and variance gridpoints. The intuition is that the observed data should reflect a specific learning process with a particular prior mean and prior variance and during the MCMC iterations I search over the prior learning parameters that maximize the likelihood. One disadvantage of the estimation approach is that the discount factor cannot be estimated and needs to be selected prior to the forward simulation procedure. However, the estimation strategy makes it relatively easy to compare a few candidate discount factors by simply adding additional parallelized forward simulations. More details on the forward simulation algorithm and bi-linear interpolation are available in the Appendix.

After forward simulation is complete the MCMC draws can proceed at usual speed. At first glance, it appears that I have simply pushed the computational intractability to the front of the estimation process, but it is important to note that each forward simulation for each grid point and each starting state can be run at the same time. With enough computers the whole procedure can be completed in minutes due to its massively parallel nature. Once the forward simulation is complete the discounted expected values are simply plugged into the likelihood and Bayesian estimation proceeds as usual.

1.5.2 Play Style Estimation

Since a portion of the gambling cost is within the player's control (average bet, decisions per hour, and time, or "BDH") I account for potential adjustments in a gambler's play style due during the forward simulation. For instance, if a gambler believes that the house advantage is very high they may decrease their next trip's average bet to reduce expected gambling costs. To estimate these effects, I estimate a regression of the log ratio of the next return trip's BDH relative to the previous trip's BDH:

$$\begin{aligned}\ln\left(\frac{\mathrm{BDH}_{i,t+1}}{\mathrm{BDH}_{i,t}}\right) &= \beta_1 A_{it} + \beta_2 \sigma_{it}^2 + \beta_3 g_{it} + \beta_4 g_{i,t+1} \\ &+ \beta_5 r_{it} + \beta_6 r_{i,t+1} \\ &+ \beta_7 o_{it} + \beta_0 + \varepsilon_{it}\end{aligned}$$

where g is the offer gaming value, r is the offer room value, and o is the outcome, represented as the casino's revenue from the player (positive values indicate a player loss), and$\varepsilon_{it} \sim \mathcal{N}\left(0, \sigma_{\varepsilon}^2\right)$. The coefficients on promotional credits and comp values control for any changes in play behavior attributed to reductions in the overall trip cost. For instance, if a player is returning on a free room offer they may increase their BDH. It is important to note that these coefficients are in regards to the *play style*, not the return decision. For example, if the coefficient on house advantage is positive it simply means that *when* the player returns they tend to play more aggressively – it does not imply that higher gambling costs increase the utility of returning to the casino.

For each of the 150 prior mean and prior variance gridpoint combinations, I run 10,000 MCMC iterations (keeping only every 10th draw) and save the posterior means. The posterior means are used during the forward simulation for adjusting BDH values as more experience signals are realized. The priors are specified as follows:

$$\begin{aligned}\beta &\sim \mathcal{N}\left(\bar{\beta}, \sigma_{\varepsilon}^{2} \cdot A^{-1}\right) \\ \sigma_{\varepsilon}^{2} &\sim \left(\nu \cdot \mathrm{ssq}\right)/\chi_{\nu}^{2}\end{aligned}$$

The coefficient estimates and prior settings are available in the Appendix.

1.5.3 Policy Function

In this section I outline the policy function used in forward simulation. The policy function estimates the probability of return given the current state. I use a Bayesian non-parametric method as outlined by Rossi (2014) to estimate this policy function. Non-parametrically, a regression models the conditional distribution of y given x. A fully non-parametric approach to regression uses the entire conditional distribution of y given x as the object of interest for inference. For the policy function I model the joint distribution of y and x and then use this joint distribution to compute the conditional distribution of $y|x$. This approach does not require assumptions and specific functional forms for how the x variables influence the conditional distribution of y.

For the policy function I estimate a five component mixture model. The covariates in x are the posterior mean and variance, predicted next trip BDH, the gambler's weeks since the last trip, month, and the room and slot promotional credit values if an offer is available during that week. I first approximate the joint distribution and then use these draws to compute the implied conditional distribution. Formally, for the rth draw with K mixing components:

$$\begin{gathered} f\left(y,x\right)^{r} = \textstyle\sum_{k=1}^{K} \pi_{k}^{r} \phi\left(y,x|\mu_{k}^{r},\Sigma_{k}^{r}\right) \\ f\left(y|x\right)^{r} = \frac{\sum_{k=1}^{K} \pi_{k}^{r} \phi\left(y,x|\mu_{k}^{r},\Sigma_{k}^{r}\right)}{f(x)^{r}} \\ f\left(x\right)^{r} = \textstyle\int f\left(y,x\right)^{r} dy = \sum_{k=1}^{K} \pi_{k}^{r} \bar{\phi}_{k}\left(x\right)^{r} \\ \bar{\phi}_{k}\left(x\right)^{r} = \textstyle\int \phi\left(y,x|\mu_{k}^{r},\Sigma_{k}^{r}\right) dy \end{gathered}$$

I use a finite mixture of normals model to simulate from the joint posterior density. The mixture of normals model is written as follows:

$$y_i \sim N\left(\mu_{\mathrm{ind}i}, \Sigma_{\mathrm{ind}i}\right)$$

$$\mathrm{ind}_i \sim \mathrm{Multinomial}\left(\pi\right)$$

Here y_i is a two dimensional vector and π is a vector of K mixture probabilities. Priors for the model are specified in conditionally conjugate forms:

$$\pi \sim \mathrm{Dirichlet}\left(\alpha\right)$$

$$\mu_k \sim N\left(\bar{\mu}, \Sigma_k \otimes a_{\mu}^{-1}\right), \quad k = 1, \ldots, K$$

$$\Sigma_k \sim IW\left(\nu, V\right)$$

Any functional of the conditional distribution such as the conditional mean can be computed based on the rth draw of the joint distribution. In the policy function, I use the conditional mean in the policy regression. The linear structure of the mixture of normals model can be exploited to facilitate computation of the conditional mean. More details are available in the Appendix.

$$\begin{aligned}
\mathbb{E}\left[y|x\right] &= \int yf\left(y|x\right)dy = \int y\frac{\sum_k \pi_k \phi\left(y,x|\mu_k,\Sigma_k\right)}{f\left(x\right)}dy \\
&= \frac{1}{f\left(x\right)}\int y\sum_{k=1}^{K}\pi_k\phi_k\left(y,x\right)dy \\
&= \frac{1}{f\left(x\right)}\sum_{k=1}^{K}\pi_k\int y\phi_k\left(y,x\right)dy \\
&= \frac{1}{f\left(x\right)}\sum_{k=1}^{K}\pi_k\int y\frac{\phi_k\left(y,x\right)}{\bar{\phi}_k\left(x\right)}\bar{\phi}_k\left(x\right)dy \\
&= \frac{1}{f\left(x\right)}\sum_{k=1}^{K}\pi_k\mathbb{E}_k\left[y|x\right]
\end{aligned}$$

1.5.4 Estimating the Value Function

Suppose $\sigma(s,\varepsilon)$ is the optimal action given state s and shock ε based on the policy function estimated in the previous section. Following Bajari et al. (2007), I take advantage of the fact that for a given learning process prior mean and prior variance, the parameters enter the utility linearly.

$$\begin{aligned}
\mathbb{E}\left[u\left(a=1,s,\varepsilon_1,\theta\right)\right] &= \Big[\left(\mathrm{BDH}\cdot A_{it}\right) \quad \left(\mathrm{BDH}\cdot A_{it}\right)^2 \quad \left(\mathrm{BDH}^2\cdot\left(\sigma_{it}^2+\sigma_{\eta_i}^2\right)\right) \\
&\qquad \text{Gaming Value}_{it} \quad \text{Room Value}_{it} \quad f\left(w_{it}\right) \quad \text{Month}_{it} \quad 1\Big]\cdot \\
&\quad \left[\theta_{1i} \quad \theta_{1i}r_i \quad \theta_{2i} \quad \theta_{3i} \quad \Omega \quad \Gamma \quad \theta_{0i}\right]' + \varepsilon_1 \\
&\equiv \Psi_{it}\cdot\left[\theta_{1i} \quad \theta_{1i}r_i \quad \theta_{2i} \quad \theta_{3i} \quad \Omega \quad \Gamma \quad \theta_{0i}\right]' + \varepsilon_1
\end{aligned}$$

Defining:

$$W\left(s;\sigma\left(s,\varepsilon\right)\right)=\mathbb{E}\left[\sum_{t=0}^{\infty}\beta^t\left(\sigma\left(s_t,\varepsilon_t\right)\Psi_{it}\right)|s_0=s\right]$$

I then have

$$V\left(s,\sigma;\theta\right)=W\left(s;\sigma\left(s,\varepsilon\right)\right)\cdot\theta$$

Exploiting this allows us to forward simulate the data only once (for each prior mean and variance gridpoint). This eases the computational burden significantly, allowing us to use the stored values when searching over the θ parameters during MCMC draws.

1.5.5 Forward Simulation & Parallelization

By taking advantage of the massively parallel structure of the forward simulation the expected value terms can be computed in a manner of hours with a reasonably sized dataset so long as many processors are available. With recent advances in online computing, estimating this complex model becomes a relatively inexpensive and fast process. To execute the forward simulation process, I use Amazon's EC2 service which rents processors at an hourly rate.[6] With small memory loads the cost is very low (less than a penny per hour) so running hundreds of instances simultaneously for a few hours is quite inexpensive.

I simply use each record in the data directly as starting states because creating starting states intended to "cover" the state space of the data is just as complex and would also require interpolation. Note that each record represents one decision period (one week), so there are hundreds of records per gambler. To give some context as to the scale of the parallelization in the empirical estimation, I conduct 100 forward simulations for each decision at each of the 287,205 rows of data over each of the 150 prior mean and variance gridpoints. This implies that theoretically the process can be divided across 8,616,150,000 servers and completed nearly instantaneously. In reality I divide the process over 30,000 servers and the process is done in about 50 hours (the servers are not all initiated at once).

At the end of the forward simulation, I obtain the expected values for returning or not at each record for each learning process prior mean and prior variance gridpoint. To recover the structural parameters, the three dimensional array (rows of data x discounted basis functions x 150 gridpoints) is then referenced during the MCMC process. I allow proposed prior mean and prior learning variances to take on any value within the range of gridpoints and use bi-linear interpolation to estimate the missing expected value. See the Appendix for more details on the bi-linear interpolation process.

1.5.6 Recap of the Estimation Procedure

For clarity, in this section I summarize the estimation procedure. First, I estimate the play style regression coefficients and policy function mixture components for each of the 150 prior mean and variance learning process parameter combinations. I do this because each prior mean and variance determines the evolution of the Bayesian updating process that each player experiences. Next, at each starting state and for each of the 150 learning

[6] https://aws.amazon.com/

process prior gridpoints I forward simulate using the play style regression coefficients and policy function parameters specific to that gridpoint. This process mimics a gambler projecting potential outcomes and then discounting the values that results in decision to return or not in a particular week. Since this process can be run in parallel across starting states and learning prior grid I divide the estimation over many cloud computers using Amazon EC2. Once the average discounted values are obtained for each of the 150 gridpoints and each record of the data, I then use standard Bayesian MCMC methods to estimate the structural parameters. As previously noted, the coefficients simply scale the values obtained from the forward simulation and because of this it is easy to introduce a hierarchical structure. The MCMC routine then searches for the structural parameters that make the observed data most likely. More details on the entire estimation procedure are available in the Appendix.

1.6 Identification

The structural parameters of interest are $\{\theta_{0i}, \theta_{1i}, \theta_{2i}, r_i, \Omega, \Gamma\}$ and the priors on each individual's learning process $\{A_{0i}, \sigma^2_{0i}\}$. Recall the choice specific value functions are as follows:

$$
\begin{aligned}
EV\left(s, a, \varepsilon_a, \theta\right) &= \mathbb{E}\left[u\left(s, a, \varepsilon_a, \theta\right)\right] + \beta \mathbb{E}_{s', \varepsilon'}\left[V\left(s', \varepsilon', \theta\right)|s, a\right] \\
EV\left(s, a=1, \varepsilon_1, \theta\right) &= \theta_{1i}\left(\text{BDH}_{it} \cdot A_{it}\right) + \theta_{1i} r_i \left(\text{BDH}_{it} \cdot A_{it}\right)^2 + \theta_{1i} r_i \left(\text{BDH}^2_{it} \cdot \left(\sigma^2_{it} + \sigma^2_{\eta}\right)\right) \\
&\quad + \theta_{2i} \text{Offer Gaming Value}_{it} + \theta_{3i} \text{Offer Room Value} \\
&\quad + \Omega f\left(w_{it}\right) + \Gamma \text{Month}_{it} + \theta_{0i} + \varepsilon_{1it} \\
&\quad + \beta \mathbb{E}\left[V\left(s', \varepsilon', \theta\right)|s, 1\right] \\
EV\left(s, a=0, \varepsilon_0, \theta\right) &= \varepsilon_{0it} + \beta \mathbb{E}\left[V\left(s', \varepsilon', \theta\right)|s, 0\right]
\end{aligned}
$$

Suppose that gamblers had complete information about the casino's house advantage.

This would imply that $A_{it} = A_i$ and $\sigma^2_{it} = 0$, and results in us being unable to separately identify A_{it} and θ_{1i}. Since gamblers observe the variation in the hold percentages, σ^2_η does not need to be estimated, unlike in Erdem and Keane (1996). Because the variability in hold percentages changes over time, it appears I can identify r_i. But since I cannot identify θ_{1i} in this complete information scenario, only the product $\theta_{1i} r_i$ is identified. So identification rests upon the assumption that incomplete information exists (which is true for a static model as well).

With incomplete information, the gambler's priors and their hold percentage exposures will guide the learning process path. Identifying the prior mean separately from the prior variance is challenging in most applications, the common solution being to fix the prior variance at one and estimate the signal variance and prior mean. But since I observe the signal variance I use the functional form of the Bayesian learning process to enable identification. A similar argument is made in Sriram et al. (2015). The priors determines how A_{it} and σ^2_{it} evolve. Thus these parameters are pinned down by the extent to which new hold percentage signals change the probability of returning (and hence the actual returns observed in the data). The hold percentage exposures vary across gamblers and create variation in the evolution of A_{it} and σ^2_{it}. So even if every gambler started with the same learning priors, the variability in outcomes across gamblers allows us to identify r_i.

1.7 Results

The results are estimated on a random subsample of 1,000 gamblers. For each gambler, 100 paths were forward simulated to derive the discounted values.[7] To assist with parameter convergence in the hierarchical model, I first estimate a homogeneous model and use those parameters as the starting values in the hierarchical estimation. The parameters are estimated using a random-walk step on each MCMC draw. Since the parameter space is quite large, I partition the estimation into four parameter blocks to make the

[7] The discounted values began to converge after averaging 50 forward simulated paths. I selected 100 to ensure consistency in the estimates.

parameter search easier (Chib and Greenberg, 1996). The first block contains the learning process prior mean and variance $\{A_0, \sigma_0^2\}$, the second contains the cost, risk aversion, offer coefficients, and intercept$\{\theta_0, \theta_1, \theta_2, \theta_3, r\}$, the third block are the coefficients on the weeks since last trip polynomials $\{\Omega\}$, and the fourth block are the month control variables $\{\Gamma\}$. Details on the estimation procedure is available in the Appendix.

In the homogeneous model, I run 80,000 MCMC draws. I discard the first 60,000 draws and keep only every 10th draw thereafter. I initialized the chain using MLE estimates. The acceptance rates of each of the four blocks is between 15% and 50% and the likelihood is -15,950. Table 2 contains the posterior means of the kept draws. As expected the coefficient on the gambling expense is negative.

Table 2: Homogeneous Results

Coefficient	Posterior Mean	SE	Coefficient	Posterior Mean	SE
Intercept	-2.8563	(1.69e-03)	γ_1	-0.3543	(1.88e-03)
A_0	0.3973	(1.54e-04)	γ_2	-0.1600	(1.82e-03)
σ_0^2	0.0012	(2.89e-06)	γ_3	-0.1878	(1.73e-03)
Cost	-0.0018	(1.72e-06)	γ_4	-0.6775	(1.88e-03)
Risk	-1.85e-07	(1.12e-09)	γ_5	-0.0153	(1.94e-03)
Gaming Offer	0.0050	(5.46e-06)	γ_6	0.0484	(1.77e-03)
Room Offer	0.0168	(5.03e-06)	γ_7	0.0775	(1.77e-03)
ω_1	0.0372	(2.54e-05)	γ_8	0.2299	(1.78e-03)
ω_2	-0.0013	(2.20e-07)	γ_9	0.2059	(1.76e-03)
ω_3	1.31e-05	(5.78e-10)	γ_{10}	-0.1051	(1.80e-03)
ω_4	-5.10e-08	(3.25e-15)	γ_{11}	-0.0606	(1.78e-03)
ω_5	6.78e-11	(2.68e-17)			

The homogeneous results are used as starting parameters for the hierarchical model. In the hierarchical model, I allow the learning process prior mean and variance, intercept, cost, risk coefficient, and offer coefficients to be a function of individual level information. The coefficients on the weeks since last trip polynomials $\{\Omega\}$and the month control variables $\{\Gamma\}$ remain fixed across the gamblers. The individual level covariates are the gambler's age, sex, distance to the casino, and an indicator for whether the gambler is at the "Gold" loyalty card status. I run 80,000 MCMC draws, discarding the first 60,000 and keeping every 10th draw thereafter. The model's likelihood is -11,358. This is a significant improvement over the homogeneous model and also greater than the likelihood from the same model with no forward looking (-11,401 when the discount factor $\beta = 0$ versus $\beta = .98$). Details on other model parameters are available in the Appendix.

Table 3 displays the estimates for the hierarchical parameters. Recall that each individual level variable influences the coefficient estimate through a multivariate regression. The individual-level variables are demeaned so that the regression intercepts reflect an "average" gambler.

Table 3: Hierarchical Interactions

		Demographics				
Coef.	Description	Int.	Age (/10)	Male	Log(miles)	Gold LP
A_0	Prior mean	0.523*	-0.014	-0.004	-0.002	-0.003
σ_0^2	Prior uncertainty	0.040*	0.000	0.001	0.001	-0.005
θ_1	Cost	-0.067*	-0.011	0.012	-0.008	-0.035
r	Risk	-0.004	0.001	-0.002	0.002	-0.002
θ_2	Offer promo credits	0.151*	0.024	0.004	0.012	0.118
θ_3	Offer room value	0.075*	-0.048	0.011	-0.141*	0.008
θ_0	Intercept	-2.793*	0.064	0.189*	-0.092*	0.287*

* = 95% highest posterior density does not cover zero

The average gambler believes that that house advantage is around 52%. While higher than the true house advantage (about 12%) gamblers have substantial uncertainty surrounding this belief, with a standard deviation of .2. As expected, the cost coefficient is negative – high house advantage perceptions lower the probability of returning. The average gambler is risk seeking (at least directionally) and the offer values significantly influence the probability of returning. The interactions with the intercept are intuitive: gamblers that live far away are less likely to return while those in the higher tier LP are more likely to return. The posterior means for all of the parameters are presented in Table 4. The results for the fixed parameters are similar to the homogeneous results.

Table 4: Full Hierarchical Results

Coefficient	Description	Posterior Mean	SE
A_0	Prior mean	0.5231	(2.40e-04)
σ_0^2	Prior uncertainty	0.0398	(2.30e-05)
θ_1	Cost	-0.0668	(1.24e-04)
r	Risk	-3.76e-03	(7.84e-05)
θ_2	Offer promo credits	0.1510	(3.88e-04)
θ_3	Offer room value	0.0765	(3.05e-04)
θ_0	Intercept	-2.7933	(4.98e-04)
ω_1	Weeks since last trip^1	0.0489	(1.82e-05)
ω_2	Weeks since last trip^2	-0.0013	(1.57e-07)
ω_3	Weeks since last trip^3	1.31e-05	(1.22e-10)
ω_4	Weeks since last trip^4	-5.10e-08	(1.59e-15)
ω_5	Weeks since last trip^5	6.78e-11	(1.22e-17)
γ_1	Jan	-0.6550	(1.90e-03)
γ_2	Feb	-0.4129	(1.72e-03)
γ_3	Mar	-0.5658	(1.75e-03)
γ_4	Apr	-0.7259	(1.85e-03)
γ_5	May	-0.2667	(1.75e-03)
γ_6	Jun	-0.1619	(1.58e-03)
γ_7	Jul	-0.1741	(1.76e-03)
γ_8	Aug	-0.0560	(1.63e-03)
γ_9	Sep	-0.0990	(1.70e-03)
γ_{10}	Oct	-0.2499	(1.77e-03)
γ_{11}	Nov	-0.1716	(1.63e-03)

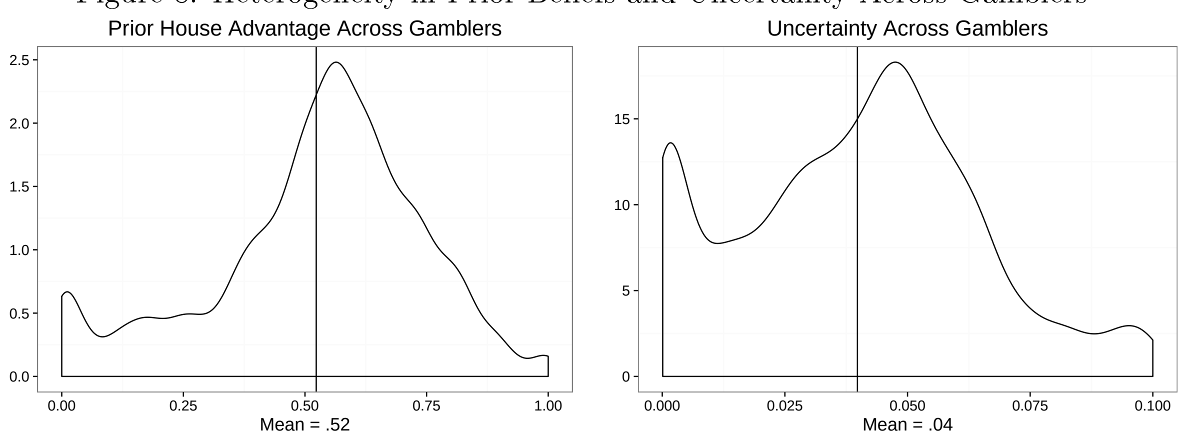
Table 5 displays the variances and correlations across the individual-level coefficient estimates. There is substantial heterogeneity across gamblers' coefficient estimates. Interestingly, there is a positive correlation between the prior mean and uncertainty: gamblers whose prior beliefs are higher tend to be more certain in their beliefs. There is also a strong negative correlation between the cost coefficient and the risk aversion; Gamblers who are more sensitive to the cost of gambling are more risk averse while those that are not as sensitive tend to be more risk seeking.

Table 5: Heterogeneity Across Gamblers

Coefficient	Description	Variance (diagonal) and Correlation (off-diagonal)						
A_0	Prior mean	.1337						
σ_0^2	Prior uncertainty	-.24	.0015					
θ_1	Cost	.06	-.03	.0563				
r	Risk	-.01	.02	-.19	.0142			
θ_2	Offer promo credits	.00	-.06	-.05	-.04	.7150		
θ_3	Offer room value	-.01	.05	.02	.05	-.55	.8034	
θ_0	Intercept	.13	-.10	.00	-.02	.06	-.06	1.0623

Figure 8 shows the distribution in posterior means across gamblers in the estimated prior house advantage and its uncertainty. Most players tend to overestimate the house advantage but the distribution is quite dispersed across gamblers. The level of uncertainty is somewhat bi-modal: while there is some mass around low uncertainty estimates there is also substantial mass around .05.

Figure 8: Heterogeneity in Prior Beliefs and Uncertainty Across Gamblers



1.8 Policy Simulations

The structural parameters are used to simulate six counterfactuals. The first two counterfactuals illustrate how projected casino revenues are quite sensitive to gamblers' prior beliefs in the house advantage and the volatility of outcomes. While these counterfactuals are informative, they do not provide casino marketers with practical solutions to act upon, for reasons to be discussed. The third and fourth counterfactuals focus on marketing solutions and show that sophisticated targeting strategies should consider how both the outcome sequence and prior beliefs may dictate where targeting is most effective. The remaining two counterfactuals explore belief-based targeting in more depth. The fifth counterfactual uses the model to identify the gamblers that are most responsive to marketing. Finally, the sixth counterfactual does a partial search for an optimal marketing strategy. While a full search is incredibly complex, the partial search still highlights that offer values should vary depending on both the outcome sequence and gambler beliefs.

1.8.1 Counterfactual 1: Accurate Prior Beliefs

In this data the average slot machine house advantage is 12.5%. The estimation results therefore suggest that gamblers overestimate the house advantage by a factor of about four prior to their first trip to the casino. Given that the cost coefficient θ_1 is negative, gamblers may be overestimating the cost of a return trip which in turn delays the return time. This counterfactual simulates expected gaming revenues under the assumption that each gambler's prior belief in the house advantage is accurate. That is, their prior belief equals the true house advantage. The results are shown in Table 6.

As expected, gamblers return at a faster rate if their prior beliefs in the house advantage are lower. With lower cost expectations gamblers no longer need many trips for their beliefs to converge to the true house advantage. Even though gamblers play less on each return trip the impact on the aggregate expected casino revenue is still positive.

If accurate beliefs in the house advantage can potentially increase long term casino revenue, why do not casino marketers simply advertise the accurate house advantages through direct mail? The primary reason is that this is not practical. Casinos tend to be very cautious on how they advertise slot machines in their direct mail offers. There is a risk that a gambler will interpret the true house advantage as a guaranteed loss limit. The casino may face backlash from the gamblers who lose more than the house advantage suggests they should. The purpose of presenting this counterfactual is to simply highlight that changes in a gambler's beliefs can have drastic long term consequences on casino revenues.

Table 6: Accurate Priors Increase Casino Revenue

	Current Prior	Accurate Prior
A_0	.523	.125
Trips	2,379	10,027
Average Weeks to Next Trip	24	16
Average Trip Slot Theoretical Loss	\$460	\$162
Total Theoretical Loss	\$1,094,114	\$1,621,907
Increase in Gaming Revenue	48.2%	
# of Gamblers Simulated	1,000	
Years Simulated	5	

1.8.2 Counterfactual 2: Slot Machine Volatility

Next, I consider the impact of reducing the volatility of the slot machine hold variance. When a casino orders a slot machine from a manufacturer they specify the variability in that machine's outcomes. In this dataset, the slot machine hold is 13.9% and has a variance of .05, meaning 98% of the hold percentages (at the trip level) are between -38% and +66%. I simulate 1,000 gamblers over 5 years to measure the revenue impact of lowering and raising the hold variance relative to its current level. The results are presented in Table 7. Figure 9 plots the casino theoretical win against a multiplier on the hold variance – the dashed line at 1 means variance is at its current level.

The simulation results show that as the volatility decreases the projected casino win increases. However, when the volatility shrinks to a point that gambler wins become very infrequent the theoretical win declines. Clearly, the volatility in the outcomes has dramatic impacts on long term casino revenues. As with the first counterfactual, even though these findings are informative they do not point to any reasonable short term solution for managers. In order for a casino to change their aggregate slot machine volatility they would need to order new slot machines and spend time installing the machines on the gaming floor. These machine, labor, and additional opportunity costs are substantial and not accounted for here.

Figure 9: Impact of Hold Percentage Volatility on Expected Casino Revenue

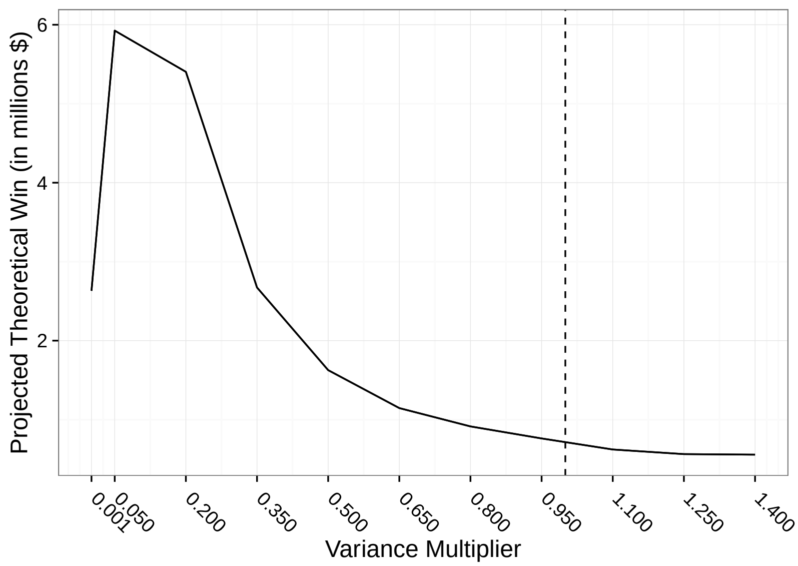


Table 7: Hold Percentage Volatility Impacts Casino Revenues

Var. Factor	Var.	1% LB	99% UB	Weeks to Return	Theo.	Total Theo.
0.001	.0002	.09	.16	29	$362	$2,630,910
0.005	.0010	.05	.20	26	$353	$2,861,178
0.010	.0021	.02	.23	23	$345	$3,138,954
0.025	0.01	-.04	.30	16	$323	$4,193,657
0.050	0.01	-.11	.37	11	$299	$5,925,798
0.200	0.04	-.35	.60	10	$319	$5,403,860
0.350	0.07	-.50	.76	15	$374	$2,672,789
0.500	0.10	-.62	.88	19	$410	$1,626,595
0.650	0.14	-.73	.99	23	$442	$1,145,819
0.800	0.17	-.82	1.08	24	$446	$913,859
0.950	0.2	-.91	1.17	23	$459	$761,572
1.100	0.23	-.99	1.24	28	$476	$621,727
1.250	0.26	-1.06	1.32	25	$465	$563,224
1.400	0.29	-1.13	1.39	26	$481	$556,337

1.8.3 Counterfactual 3: Incorporating Gambler Outcomes with Naive Targeting

The first two counterfactuals illustrate that changes in prior beliefs and hold percentage volatility can have substantial impacts on long term casino revenue. However, as discussed the results alone do not lend themselves immediately to practical solutions for managers. The purpose of these remaining four counterfactuals is to show how targeted marketing could be used in conjunction with the outcomes and player beliefs to improve casino profitability.

In this counterfactual I compare three marketing strategies: 1) the industry standard of basing offer values on gamblers' theoretical losses ("Industry"), 2) basing offer values on actual outcomes but excluding gamblers players who won on their last trip ("Actual ex Wins"), and 3) basing offer values on theoretical losses (similar to the industry standard) but again excluding gamblers who won on their last trip ("Theo ex Wins"). The second and third strategies are meant to represent naive targeting strategies: gamblers who win are more likely to have low beliefs in the house advantage and therefore should be more likely to return to the casino anyways. Given this, the casino may be able to save on marketing expenses by excluding these players from offers. Furthermore, in the second scenario the casino provides an incentive to return that is directly in line with the loss experienced. I consider these strategies "naive" because they do not consider how each gambler's beliefs in the house advantage may influence the effectiveness of marketing – only the outcomes are used.

As in the empirical data, in each decision period there is a 45% chance that the gambler will be exposed to a marketing offer. The offers are valued at 30% of their last trip's theoretical or actual win. The total offer value is split into a room component and promotional credits, with two-thirds of the total offer value going to the room and one-third going to promotional credits. In the industry standard simulation, all gamblers have an opportunity to obtain an offer but in the "Actual ex Win" and "Theo ex Win" simulations offers will not be available to gamblers who won on their last trip.

The results in Table 8 show that both naive approaches to targeting are less profitable than the current industry standard. In the "Actual ex Wins" scenario, top line revenue remains relatively constant but the overall promotional costs are higher, even though fewer offers were redeemed. This may seem counterintuitive but this is because actual outcomes tend to have much more variability relative to theoretical outcomes, especially when evaluated at the trip level (as data becomes aggregated the theoretical outcomes converge to actual outcomes). In the "Theo ex Wins" strategy, promotional costs decrease dramatically but top line revenue also suffers. The short term gains that might be had from the reduction in promotional costs is offset by longer intervisit times. The results suggest that strategies that appear intuitive at first are not always more profitable in the long term. This counterfactual emphasizes the need for a more sophisticated targeting strategy.

Table 8: Naive Targeting is Ineffective

Targeting Criteria	Industry	Actual ex Wins	Theo. ex Wins
Trips	3,084	2,441	2,388
Avg. Weeks to Return	19	25	24
Avg. Theoretical Win/Trip	$472	$498	$482
Total Theoretical Win	$1,456,769	$1,215,938	$1,150,119
Total Actual Win	$1,453,383	$1,408,149	$1,266,449
Promotions Redeemed	1,328	724	697
Room Value	$128,509	$168,937	$65,902
Promotional Credits	$64,254	$84,469	$32,951
Room Cost ($30/night)	$43,380	$41,460	$21,840
Promo Credit Cost (1 cycle)*	$56,222	$73,910	$28,832
Net Theoretical Win	$1,357,167	$1,100,568	$1,099,446
Net Actual Win	$1,353,780	$1,292,779	$1,215,777

*The cost of promotional credits is not a certainty since wins can be cycled back into the machine and generate additional payouts. See the Appendix for a discussion.

1.8.4 Counterfactual 4: Marketing Impact by Past Outcomes and Beliefs

This simulation extends the previous by incorporating prior beliefs into the targeting decision. Table 9 shows the impact of marketing when gamblers' prior beliefs and uncertainty are high or low and when gamblers are either winning or losing. The impact of marketing is measured by comparing overall expected casino revenue with marketing versus without marketing. For example, an impact of .1 means that there is a 10% increase in revenue across gamblers in the presence of marketing. The marketing rule imposed is the same as the industry standard as described in the previous counterfactual. The "high" and "low" categorizations are set using the 5th and 95th percentile estimated prior means and prior variances.

When a gambler's prior belief in the house advantage is high and their uncertainty is high, marketing is more impactful if the player is on a *winning* streak rather than a losing streak. However, for gamblers whose prior beliefs in the house advantage are low, marketing is more impactful when players are on a *losing* streak. Notice that marketing is ineffective for gamblers whose beliefs in the house advantage are very high and their uncertainty is very low. This is intuitive: these gamblers are very certain that the cost is very high and because of they will not return regardless of marketing offers.

This simulation emphasizes the importance of considering both the prior beliefs *and* the outcome sequence when designing the targeting strategy. In the previous counterfactual the naive assumption was that only the outcome mattered but here I see that it is a combination of the outcomes and prior beliefs that dictate where marketing is more effective. This insight is very useful to managers who need to allocate their limited marketing budget across gamblers.

Table 9: Marketing Impact Depends on Prior Beliefs

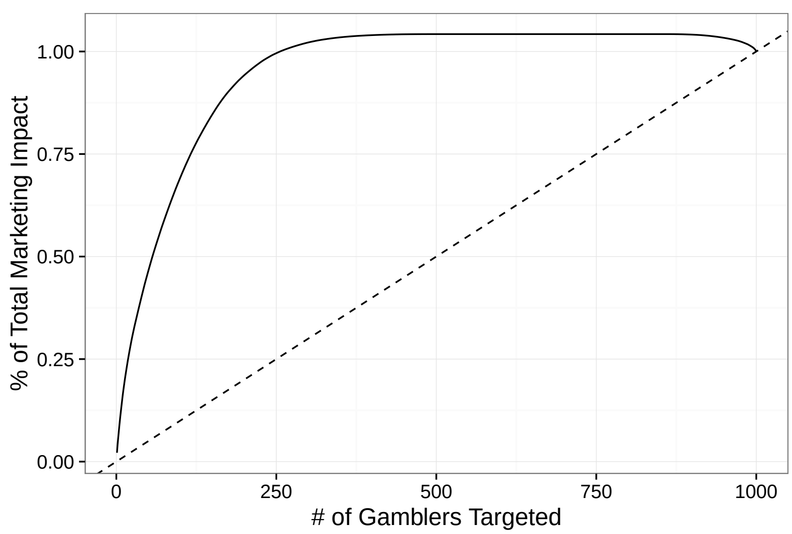
Prior Belief in House Advantage	Prior Uncertainty	Player Winning	Player Losing	Δ
High	High	3.8	1.3	2.6
High	Low	0.0	0.0	0.0
Low	High	0.1	3.4	-3.3
Low	Low	0.3	2.4	-2.1

1.8.5 Counterfactual 5: Marketing Impact by Gambler

In this counterfactual I analyze the relationship between gamblers' *posterior* beliefs and the marketing impact. The posterior beliefs summarize both the prior beliefs and the outcome sequences realized, thereby reducing the number of metrics managers need to consider for targeting. To add more realism to this simulation, I use the 1,000 gamblers from the dataset rather than creating artificial gamblers. I simulate five years worth of gambling activity, picking up where the observed data ends. Again the focus is on the impact of marketing, meaning the change in expected casino revenue when there is marketing versus no marketing present. The goal of this simulation is to identify gamblers where marketing has the greatest impact and then determine if the marketing impact is in any way related to posterior mean and uncertainty in the belief of the house advantage.

Figure 10 shows the marketing impact represented by a lift chart. If gamblers were randomly targeted the total impact is expected to follow the dashed line. However, the simulations allow us to identify the gamblers where marketing will likely have the greatest impact.[8] Notice that just about all of the gains from marketing activity are realized from about one quarter of the gamblers. The other gamblers are not impacted by the marketing activity or in a few rare cases the marketing actions actually result in declines in gaming revenue.

Figure 10: Targeted Marketing Lift



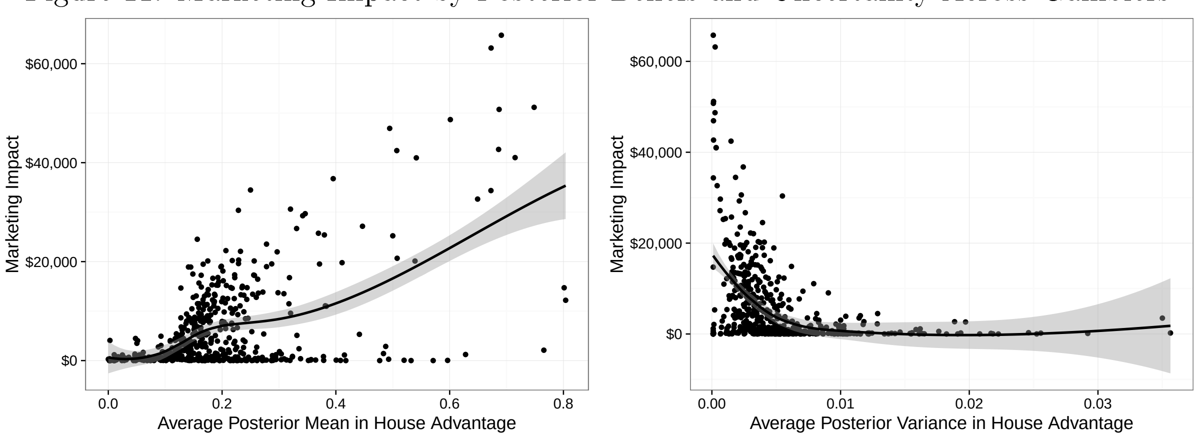
For gamblers who are most impacted by marketing (those in the front of the curve where the cumulative impact is less than 99%), the posterior belief in the house advantage tends to be higher and the uncertainty much lower.

[8] Since the impact of marketing depends on the outcome sequence a more thorough analysis would simulate over many potential outcome paths. I conducted a simulation setting the hold percentage to a constant (the mean) and the interpretations are the same.

	# of Gamblers	Post. Mean	Post. Uncertainty
99% of Marketing Impact	242	.227	.0036
Remaining 1%	758	.196	.0076

Figure 11 illustrates the differences across gamblers. For each of the 1,000 gamblers, the marketing impact is plotted against the posterior mean and posterior uncertainty averaged across all of their realized return trips. Marketing has a greater impact on gamblers with higher posterior means and lower uncertainty. The correlation between the posterior mean and posterior variance across gamblers is -.128: gamblers with higher beliefs in the house advantage tend to have less uncertainty. While this may seem to contradict the findings from the previous counterfactual it is important to note that the previous counterfactual examined the extremes of beliefs, at the 5th and 95th percentile. In addition, this counterfactual uses the actual gamblers, rather than simulated gamblers. In both cases the fact remains that there is a strong relationship between the posterior beliefs and the impact of marketing.

Figure 11: Marketing Impact by Posterior Beliefs and Uncertainty Across Gamblers



1.8.6 Counterfactual 6: Optimal Marketing Offers

The previous counterfactual provides evidence that posterior beliefs influence the impact of marketing. A natural extension is to then search for the optimal marketing strategy. That is, for each gambler and each outcome experience which offer strategy will lead to the highest long term expected revenue? Finding the global optimum is very difficult (at least in this casino example) because each room value and slot promotional credit combination would need to be evaluated for each gambler at each decision period for every potential outcome sequence. Even though finding the global optimum is incredibly complex this counterfactual shows that even a relatively simple constrained optimization can lead to substantial improvements in projected revenue.

In this constrained search I vary the slot promotional credits and bin the posterior beliefs into four categories. The goal is to determine how much each of the four posterior belief categories should receive in slot promotional credits. In this dataset the promotional credit value is typically set at 10% of the past theoretical loss level. I simulate this baseline percentage and four alternatives: 0%, 5%, 15%, and 20%. The belief and uncertainty levels are grouped into four categories: high/low belief in the house advantage and high/low uncertainty. The cutoff for the belief in the house advantage is the casino's true house advantage and the cutoff for the uncertainty is based on a median split of the observed gambler's posterior variances.

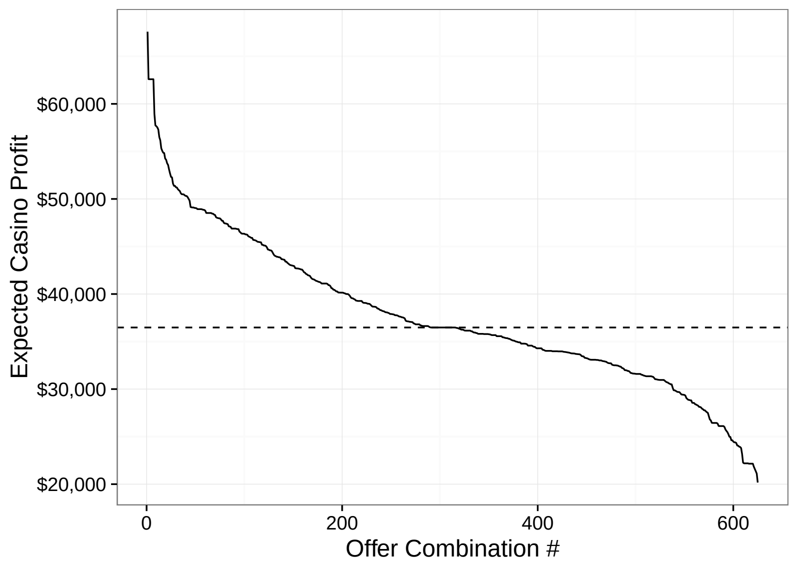
Category	Belief in House Advantage	Uncertainty in Belief
Low	$<$12.5%	$<$.0029
High	$>=$12.5%	$>=$.0029

Another challenge in searching for the optimal marketing offer is that gamblers can switch categories over time depending on their outcomes. That is, they may start in a high belief/high uncertainty state, move to high belief/low uncertainty state, and then end in a low belief/low uncertainty state. Because each state will have its own marketing strategy, all 625 combinations of offers need to be considered: five promotional credit percentages in each of the four offer states.

I simulate one hundred gamblers for two years in each of the 625 offer value combinations. Each gambler starts with the same prior beliefs and uncertainty on the house advantage (based on the hierarchical results for the "average" gambler). Profit is obtained by subtracting room and promotional credit costs from the projected casino revenue.

Figure 12 shows the sorted profit across all 625 simulations. The dashed line shows baseline profitability where the four belief categories each receive promotional credits valued at 10% of theoretical losses. The range in the profit is substantial: the top strategies generate over \$55,000 in profit while the worst strategies generate around \$25,000.

Figure 12: Simulated Casino Profit by Targeting Strategy

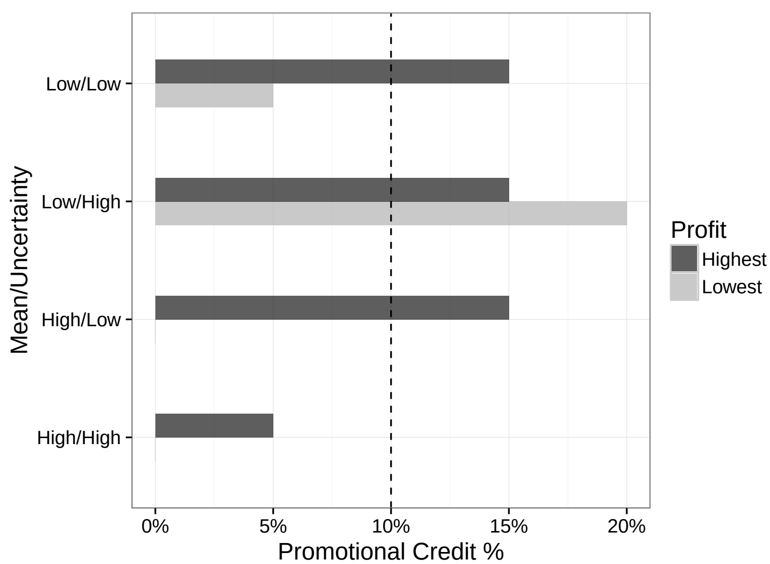


Rather than try to evaluate each of the 625 simulations individually, I instead compare the differences in the most and least profitable strategies, shown in Figure 13. This shows which promotional credit percentage is associated with the most and least profitable strategies in each of the four belief/uncertainty categories. Notice that the most profitable strategy does not use the baseline percentage of 10% in any of the four belief categories:

when the belief in the house advantage is below the actual house advantage, a higher percentage is recommended whereas when the belief in the house advantage is high the policy depends on the uncertainty. It is also interesting to note that the most profitable strategy does not max out or eliminate the promotional credit amount in any of the four belief bins, suggesting that the solution is contained within the boundaries of the simulation. The most profitable strategy generated $62,290 in profit, compared to $36,479 in the baseline scenario where all gamblers receive the same promotional credit percentage regardless of their beliefs in the house advantage, an increase of 85.3%. For a more conservative (and realistic) measure of success, the top half of the strategies still increased baseline profit by an average of 19.7%.

The model presented provides a framework for managers to use in order to target gamblers based on their beliefs and outcome sequences. This simulation shows that the gains from doing so can be significant, even when the strategy employed is the result of a heavily constrained search.

Figure 13: Constrained Optimization of Best and Worst Strategy Profiles



1.9 Discussion and Conclusion

This paper uses a dynamic Bayesian learning framework to develop a methodology that assists managers in direct marketing targeting when the customer's experiences are random and these variations are observed by the firm. This class of problem arises naturally in many industries, specially those that are service-based where across-occasion variation in experience quality can be high. Depending on the signal variability, it may take many experiences for the customer to learn the true distribution of quality. Until the true distribution is learned, the customer will likely have biased perceptions. If a customer's initial experiences are likely to lead to an inference that the quality level is lower than what it is in truth, the situation may warrant additional targeted marketing. To facilitate tractable estimation I take advantage of inexpensive cloud computing and exploit the massively parallel structure of combining forward simulation with a utility function that is linear in parameters. The proposed structure easily incorporates flexible heterogeneity distributions to generate individual-level parameter estimates, which is central to many modern targeted marketing problems.

The proposed methodology is illustrated using data from a casino where gamblers need to learn about the average slot machine house advantage. Gamblers use their beliefs on the house advantage to project future trip costs which in turn influences when they return to the casino and how they play on a return trip. The gaming industry offers an attractive setting to illustrate this methodology, for a variety of reasons, one of which is that exogenous gambling outcomes provide many distinct and unique experience sequences at the gambler level. The results and counterfactuals suggest that gamblers tend to overestimate the house advantage, which increases gambler's projected gambling expenses and reduces the probability of returning to the casino within a specific interval. The counterfactuals also highlight the importance of incorporating gambler beliefs and outcomes into the marketing decisions. The simulations suggest that campaign effects can be increased greatly when the targeting considers the effect of marketing on customer beliefs and visit decisions.

There are a few limitations worth addressing in regards to the estimation strategy and the empirical analysis. While the estimation strategy does improve the tractability substantially, given that it is not a full solution approach, it is difficult to pre-determine the number of paths to forward-simulate for each starting state to ensure the state space is sufficiently explored. In our estimation, I continued to simulate additional paths until the discounted future values appeared to converge – I then doubled this number of simulations as an extra precaution in the final estimation. Another limitation, mentioned already, is that the discount factor cannot be estimated, but again it is relatively easy to test multiple discount factors by taking advantage of the parallel design. One limitation with the empirical analysis is that competitor activity is not observed. This includes both player activity at competitor casinos and competitor marketing activity. For instance, I do not know if delays in return trips are due to gamblers visiting other casinos, the marketing activities of competitor casinos, or simply a lack of gambling. A more central concern is that gamblers may be learning from slot machines at outside casinos between the casino trips I observe. At first glance, multi-casino gambling appears to impact the learning process substantially. However, it is important to note that each casino will have its own mix of slot machines, which means the average house advantage of each casino is likely to be different. Even if a gambler visits other casinos, they still need to learn about the *focal* casino's house advantage. Finally, there are some characteristics on the counterfactuals that warrant discussion. First, I do not account for competitor reactions. However, competitor reactions are unlikely to be much of a concern because 1) competitors will not know which gamblers have been identified as being responsive to marketing, and 2) competitors do not know the outcome sequences experienced by gamblers at this casino. In other words, given the targeting strategies presented it is not immediately clear how a competitor could react. Second, I do not allow gamblers to learn about the targeting strategy over time. However, since the relatively simple targeting strategies were shown to be ineffective a gambler would need to learn a very sophisticated strategy, which is quite unlikely unless they have an unrealistically high number of trips and marketing exposures.

There are many possible extensions to this current work. One important extension is to allow a learning rate to be estimated at the individual level. In this paper the learning rate is fixed due to the formulation of the Bayesian updating process but by including additional parameters, I can capture the phenomenon that gamblers learn at different rates, similar to Narayanan and Manchanda (2009) but with forward-looking consumers. The speed of learning could have substantial impacts on the targeted marketing decisions. Another extension could account for how beliefs influence projected marketing offers, which in turn may change current-period decisions. Finally, there may be more efficient ways to search for the global optimal marketing strategy when using this model for targeting. I leave these topics for future research.

1.10 Appendix 1.A Conjugate Prior in Truncated Normal Distribution

If the prior is truncated normal and the signal is an unbounded normal then the corresponding posterior is also a truncated normal. In other words, the truncated normal distribution is also a conjugated prior for a standard normal likelihood of signal generation. This proof is similar to the one in Li (2014).

Theorem 1. Suppose the parameter of interest θ is distributed in normal distribution truncated at 0 and 1, i.e., $\theta \sim \mathcal{TN}\left(\mu_0, \sigma_0^2 = \lambda_0^{-1}, 0, 1\right)$, and the likelihood for signal

$$x = \theta + \xi$$

where $\xi \sim \mathcal{N}\left(0, \sigma_\xi^2 = \lambda_\xi^{-1}\right)$, then the posterior distribution

$$\theta | x \sim \mathcal{TN}\left(\mu_1, \sigma_1^2 = \lambda_1^{-1}, 0, 1\right)$$

with

$$\begin{aligned} \mu_1 &= \frac{\lambda_0}{\lambda_0 + \lambda_\xi}\mu_0 + \frac{\lambda_\xi}{\lambda_0 + \lambda_\xi}x \\ \lambda_1 &= \lambda_0 + \lambda_\xi \end{aligned}$$

Proof. Let $\phi\left(t, \mu, \sigma^2\right)$ be the normal pdf with mean μ and variance σ^2, and $\Phi\left(t, \mu, \sigma^2\right) = \int_{-\infty}^{t} \phi\left(s, \mu, \sigma^2\right) ds$ be the CDF. I know that

$$\begin{aligned} f(\theta) &= \frac{\phi\left(\theta, \mu_0, \sigma_0^2\right)}{\Phi\left(1, \mu_0, \sigma_0^2\right) - \Phi\left(0, \mu_0, \sigma_0^2\right)} \\ f(x|\theta) &= \phi\left(x, \theta, \sigma_\xi^2\right) \end{aligned}$$

so

$$f(x) = \int_0^1 f(x|\theta) f(\theta) \, d\theta = \frac{\int_0^1 \phi\left(x, \theta, \sigma_\xi^2\right) \phi\left(\theta, \mu_0, \sigma_0^2\right) d\theta}{\Phi\left(1, \mu_0, \sigma_0^2\right) - \Phi\left(0, \mu_0, \sigma_0^2\right)}$$

and

$$\begin{aligned} f(\theta|x) &= \frac{f(\theta) f(x|\theta)}{f(x)} = \frac{\phi\left(x, \theta, \sigma_\xi^2\right) \phi\left(\theta, \mu_0, \sigma_0^2\right)}{\int_0^1 \phi\left(x, \theta, \sigma_\xi^2\right) \phi\left(\theta, \mu_0, \sigma_0^2\right) d\theta} \\ &= \frac{\phi\left(x, \theta, \sigma_\xi^2\right) \phi\left(\theta, \mu_0, \sigma_0^2\right) / \int_{-\infty}^{\infty} \phi\left(x, \theta, \sigma_\xi^2\right) \phi\left(\theta, \mu_0, \sigma_0^2\right) d\theta}{\int_0^1 \phi\left(x, \theta, \sigma_\xi^2\right) \phi\left(\theta, \mu_0, \sigma_0^2\right) d\theta / \int_{-\infty}^{\infty} \phi\left(x, \theta, \sigma_\xi^2\right) \phi\left(\theta, \mu_0, \sigma_0^2\right) d\theta} \\ &= \frac{\phi\left(\theta, \mu_1, \sigma_1^2\right)}{\int_0^1 \phi\left(\theta, \mu_1, \sigma_1^2\right) d\theta} = \frac{\phi\left(\theta, \mu_1, \sigma_1^2\right)}{\Phi\left(1, \mu_1, \sigma_1^2\right) - \Phi\left(0, \mu_1, \sigma_1^2\right)} \end{aligned}$$

where the second equality before last is obtained by the standard conjugate prior of the normal distribution, which is

$$\begin{aligned} \theta &\sim \mathcal{N}\left(\mu_0, \sigma_0^2\right) \\ x|\theta &\sim \mathcal{N}\left(\theta, \sigma_\xi^2\right) \end{aligned}$$

will imply

$$\theta|x \sim \mathcal{N}\left(\mu_1, \sigma_1^2\right)$$

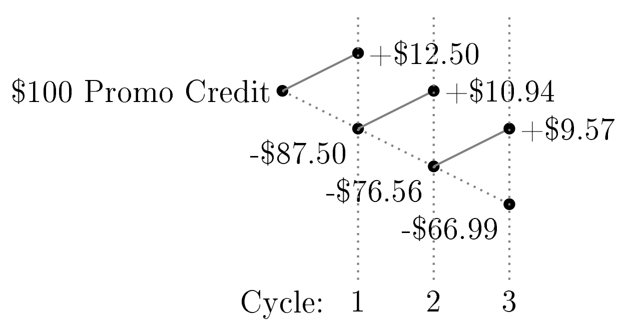
1.11 Appendix 1.B Forward Simulation Algorithm

```
Algorithm: Forward Simulation Pseudo-Code
for simulation r ← 1 to R do
    foreach starting state s_0 do
        foreach starting action a_0 ∈ {0, 1} do
            for time period t ← 0 to ∞ do
                if t_0 = 0 then
                    returnFlag = a_0
                else if t_0 > 0 then
                    Calculate probability of returning based on state
                    Draw from uniform [0,1] and update returnFlag based on
                      probability
                end
                Record the current state values (to be discounted)
                if returnFlag = 1 then
                    Reset the weeks since last trip
                    Increment the month (if needed)
                    Draw a new hold value
                    Update beliefs based on hold
                else if returnFlag = 0 then
                    Increment week counter
                    Increment the month (if needed)
                end
                Update play style for next trip
                Draw a new marketing offer
            end
            Discount the values by β
        end
    end
end
```

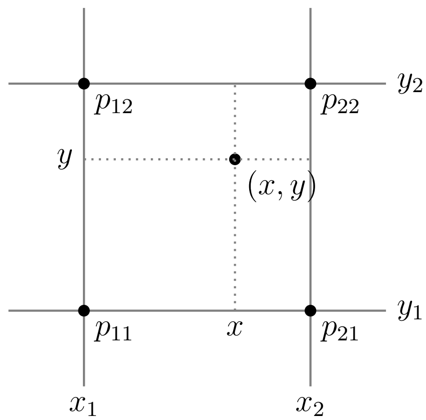
Algorithm 1: Forward Simulation

1.12 Appendix 1.C Cost of Promotional Credits

The casino industry remains divided on the cost of promotional credits. The reason is because the true cost depends on how many times the credit is cycled back through the slot machine, as illustrated below. In this example, a \$100 promotional credit is played through a slot machine with a 12.5% house advantage. That means the casino is expected to keep \$12.50 and pay out \$87.50, which the gambler can keep and convert to cash. However, if the gambler cycles their winnings back into the machine the casino is expected to keep another 12.5% (or \$10.94) and pay back \$76.56, which is now the true cost of the promotional credit. Obviously, this can continue until the gambler has cycled the all of the winnings through the machine, resulting in zero cost to the casino. It becomes even more difficult because not all slot machines have the same house advantage, and winnings from one machine can be played on another. Because of these complexities, financial planning departments often use simple rules such as charging 100% of the cost or some fixed fraction of projected promotional credits redeemed.



1.13 Appendix 1.D Bi-linear Interpolation



Bi-linear interpolation is an extension of linear interpolation. Linear interpolation is performed in one direction, and then again in the other. In this analysis, the learning process prior mean and variance gridpoints form the bi-linear lattice. The forward simulation yields discounted expected values at each prior mean and prior variance gridpoint (labeled p_{11}, p_{12}, p_{21}, and p_{22} in the figure above). During the MCMC draws, a new prior mean and prior variance may be suggested that has not been forward simulated over, in this example point (x, y). The below formula is used to approximate that point's discounted expected value, $f(x, y)$.

$$a = f(p_{11})(x_2 - x)(y_2 - y) + f(p_{21})(x - x_1)(y_2 - y)$$

$$b = f(p_{12})(x_2 - x)(y - y_1) + f(p_{22})(x - x_1)(y - y_1)$$

$$f(x, y) = \frac{a + b}{(x_2 - x_1)(y_2 - y_1)}$$

1.14 Appendix 1.E Estimation Procedure

1.14.1 Step 0. Forward Simulation Settings

The casino data is constructed so that each record represents a decision period. The first record for each gambler is the first trip to the casino and each week after the first trip is a new record. In preparation for the forward simulation, the discount factor, number of periods to forward simulate, and number of forward simulations per starting state need to be selected.

Parameter	Description	Value
β	discount factor	.98
T	time periods to forward simulate	228
R	number of forward simulations per starting state	100

I selected T such that it is the smallest value for which β^T is less than .01 so that simulating beyond this point will have negligible impacts on the discounted values. I selected R based on experimentation; I found that at around $R = 50$ the discounted values began to converge. As a conservative measure I then doubled this to $R = 100$.

The forward simulation is done across 150 prior mean and prior variance combinations. This consists of 15 prior mean values and 10 prior variance values, given below. As illustrated in the paper, the gridpoints were selected in order to represent a broad range of truncated normal distributions.

A_0 Grid	σ_0^2 Grid
0.0000010	0.0001
0.0300010	0.0041
0.0600010	0.0081
0.0900010	0.0121
0.1200010	0.0161
0.1500010	0.0201
0.1800010	0.0241
0.2100000	0.0300
0.3228571	0.0650
0.4357143	0.1000
0.5485714	
0.6614286	
0.7742857	
0.8871429	
1.0000000	

1.14.2 Step 1. Play Style and Policy Function Estimation

For each gridpoint, I estimate the coefficients for the play style regression and the non-parametric policy function. These estimates need to be done for each gridpoint because each prior mean and prior variance changes how the posterior beliefs evolve. In other words, the coefficient estimates depend on the learning prior settings. The play style priors are set to the following:

$$\bar{\beta} = 0$$
$$A = .01$$
$$\nu = 3$$
$$\text{ssq} = \text{var}(y)$$

Here is a summary the posterior means of the play style coefficient estimates across the 150 gridpoints:

	β_1	β_2	β_3	β_4	β_5	β_6	β_7	β_0
25%	1.5203	-3.3670	-0.00138	0.00151	-0.00145	0.00163	-0.00048	-0.25878
50%	2.2112	2.0343	-0.00135	0.00154	-0.00140	0.00164	-0.00048	-0.19895
75%	2.5324	13.4973	-0.00133	0.00155	-0.00137	0.00165	-0.00046	-0.14189
Min	0.0706	-8.5708	-0.00155	0.00139	-0.00159	0.00160	-0.00049	-0.36145
Median	2.2112	2.0343	-0.00135	0.00154	-0.00140	0.00164	-0.00048	-0.19895
Mean	2.0307	6.6500	-0.00137	0.00153	-0.00142	0.00164	-0.00047	-0.18996
Max	3.7892	51.9811	-0.00130	0.00157	-0.00136	0.00167	-0.00043	0.17292

Details of the non-parametric policy function estimation are given in the paper. The prior settings used are as follows:

$$\alpha = 5$$
$$\bar{\mu} = 0$$
$$a_u = .01$$
$$\nu = 8$$
$$V = \text{diag}(\nu)$$

1.14.3 Step 2. Forward Simulation

At this point the data is ready for forward simulation. I divided the data into 200 groups, each group contained the data associated with 5 gamblers. I sent each of the 200 groups with each of the 150 gridpoints to Amazon EC2, meaning a total of 30,000 files would be forward simulated over and then collected. The parallelization can be as granular as one wishes however since Amazon rounds up on each instance hour consumed I divided the data based on how many paths I could comfortably process in one hour (allowing for time to send and receive the data from Amazon via ssh). Otherwise large amounts of processing time (and cost) would be wasted. The sending and receiving of data to Amazon was done automatically through a shell script – the instances were stagger started to avoid overloading the ssh calls and data transfers.

For a single simulation (i.e., 1 of the 30,000 files) I transferred to Amazon EC2 the following: 1) the data from one group of 5 gamblers, 2) the play style and policy function coefficients associated with one gridpoint, 3) a generic c++ file to execute the forward simulation algorithm, and 4) a generic file that loads the starting data and coefficient estimates, runs the forward simulation algorithm, and saves the output. Within an hour the shell script revisited the Amazon EC2 instance and collected the output file. The 30,000 output files were then consolidated in preparations for the Bayesian estimation. Details of the forward simulation algorithm are available in the Appendix of the paper.

1.14.4 Step 3. Bayesian Estimation

After the forward simulation files were consolidated, I proceeded with the Bayesian estimation. The consolidation file is structured into an array where the rows are each row of the data, the columns are the discounted values, and the slices represent the gridpoints. The memory demands of this array are the primary reason I had to limit the estimation to 1,000 gamblers. Recall that since the parameter space is large, the Metropolis-Hastings random walk procedure is split into four partitions:

Partition	Parameters
BLOCK 1	$\{A_0, \sigma_0^2\}$
BLOCK 2	$\{\theta_0, \theta_1, \theta_2, \theta_3, r\}$
BLOCK 3	$\{\Omega\}$
BLOCK 4	$\{\Gamma\}$

Below is an outline of the homogeneous estimation procedure.

1. Initiate all parameters at "old" values.

2. For each MCMC draw:

   (a) Draw "new" BLOCK 1 learning process priors A_0 and σ_0^2 and truncate if they are outside of the grid range.

   (b) Using bi-linear interpolation, interpolate "new" discounted forward simulation values using the newly drawn learning process priors.

   (c) Draw "new" values for all other parameters: BLOCK 2, BLOCK 3, and BLOCK 4.

   (d) Process BLOCK 1 parameters:

      i. Keeping BLOCK 2, BLOCK 3, and BLOCK 4 parameters fixed at the "old" levels, compute $\alpha = \min\left[1, \frac{L^{new}}{L^{old}}\right]$ and accept the "new" BLOCK 1 parameters with probability α.

      ii. If the BLOCK 1 parameters are accepted, override the "old" values with the "new" values.

   (e) Process BLOCK 2, BLOCK 3, and BLOCK 4 parameters in a similar fashion.

3. End MCMC loop

In the homogeneous estimation, the "new" parameters are generated by a random-walk process. That is, $\theta^n = \theta^o + \varepsilon$, where $\varepsilon \sim \mathcal{N}(0, \sigma^2)$. Each parameter is drawn independently of the others. The variances of each parameter are given below:

Coefficient	Description	Random-Walk Variance
θ_0	Intercept	5e-5
A_0	Prior Mean	1e-5
σ_0^2	Prior Variance	5e-8
θ_1	Cost	1e-9
r	Risk	1e-17
θ_2	Gaming Offer	1e-7
θ_3	Room Offer	1e-7
ω_1	Weeks since last trip^1	2e-8
ω_2	Weeks since last trip^2	2e-10
ω_3	Weeks since last trip^3	4e-19
ω_4	Weeks since last trip^4	2e-28
ω_5	Weeks since last trip^5	2e-33
$\gamma_1 - \gamma_{11}$	Monthly Dummies	3e-3

For the hierarchical estimation, all variables except the $\{\Omega\}$ weeks since last trip polynomials and $\{\Gamma\}$ monthly dummy variables are estimated at the individual level. Below is an outline of the heterogeneous estimation procedure:

1. Initiate all parameters at "old" values.

2. Initiate the heterogeneity variance: Σ_θ.

3. For each MCMC draw:

   (a) Calculate the random walk step for the individual level parameters: $s^2 \times \Sigma_\theta$.

   (b) For each gambler i:

      i. Draw "new" individual level parameters $\{A_{0i}, \sigma_{0i}^2, \theta_{0i}, \theta_{1i}, \theta_{2i}, \theta_{3i}, r_i\}$ using variance Σ_θ.

      ii. Truncate A_{0i} and σ_{0i}^2 if they are outside of the grid range.

iii. Using bi-linear interpolation, interpolate "new" discounted forward simulation values using the newly drawn learning process priors.

iv. Process BLOCK 1 parameters for gambler i:

A. Keeping BLOCK 2, BLOCK 3, and BLOCK 4 parameters fixed at the "old" levels, compute $\alpha = \min\left[1, \frac{L_i^{new}}{L_i^{old}}\right]$ and accept the "new" BLOCK 1 parameters with probability α.

B. If the BLOCK 1 parameters are accepted, override the "old" values with the "new" values.

v. Process BLOCK 2 parameters for gambler i in a similar fashion to BLOCK 1.

(c) End gambler loop

(d) Draw "new" Δ and Σ_θ using a multivariate regression of $\Theta = Z\Delta$, where the ith row of Θ is gambler i's θ and the ith row of Z is gambler i's demographics.

(e) Draw "new" values for BLOCK 3 and BLOCK 4 parameters. As in the homogeneous estimation, each parameter is drawn independently of each other.

(f) Process BLOCK 3 parameters:

i. Keeping the BLOCK 4 parameters at the "old" levels and the BLOCK 1 and BLOCK 2 parameters at the mean level across gamblers, compute $\alpha = \min\left[1, \frac{L^{new}}{L^{old}}\right]$ and accept the "new" BLOCK 3 parameters with probability α.

ii. If the BLOCK 3 parameters are accepted, override the "old" values with the "new" values.

(g) Process BLOCK 4 parameters in a similar fashion to BLOCK 3.

4. End MCMC loop

The heterogeneous prior settings are given below:

$$
\bar{\Delta} = \begin{array}{cccccccc}
 & A_0 & \sigma_0^2 & \theta_1 & \theta_2 & \theta_3 & \theta_4 & \theta_0 \\
\text{Intercept} & 0.397 & 0.001 & -0.002 & 0.000 & 0.005 & 0.017 & -2.860 \\
\text{Age } (/10) & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\text{Male} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\text{Log (Distance)} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\text{Gold LP Card} & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{array}
$$

$$
A = \operatorname{diag}(1)
$$

$$
\nu = 10
$$

$$
\Sigma = \begin{bmatrix}
1e-5 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1e-8 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1e-9 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1e-17 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1e-7 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1e-7 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 5e-5
\end{bmatrix}
$$

$$
s^2 = .2
$$

The random walk step sizes for the heterogeneous BLOCK 3 and BLOCK 4 parameters are the following:

Coefficient	Description	Random-Walk Variance
ω_1	Weeks since last trip^1	1e-9
ω_2	Weeks since last trip^2	1e-11
ω_3	Weeks since last trip^3	2e-20
ω_4	Weeks since last trip^4	1e-29
ω_5	Weeks since last trip^5	1e-34
$\gamma_1 - \gamma_{11}$	Monthly Dummies	5e-3

2 Identifying Coalition Membership Benefits through Evolving Networks

2.1 Introduction

A coalition loyalty program offers incentives to customers at multiple businesses. In this type of loyalty program, stores join a coalition "network" and customers earn rewards at a higher rate when shopping at stores that are part of the network versus when shopping elsewhere. The stores in the network do not have to be related in any way and are connected only through the coalition program. Individual stores can benefit from the branding and marketing power of the coalition and customers benefit by earning rewards through a wide variety of outlets. The coalition structure is a cost-effective way to incentivize customers through a variety of shopping choices (Berman, 2006).

Coalition loyalty programs are popular in Europe and are beginning to gain popularity in the United States (e.g., Plenti) (Capizzi and Ferguson, 2005). However, prior research is still mixed on the effectiveness of coalition programs. Dorotic et al. (2011) found that cardholders use their cards for regular purchases can collect rewards, but generally do not change their purchase behavior once enrolled in the LP. Moore and Sekhon (2005) found similar results, but find that consumers are willing to alter behavior so long as the motivation is sufficient.

A key question of interest to both managers of coalition programs and potential network entrants is to determine the value of participating in the network. This research seeks to answer that question. From an academic perspective, this research furthers our understanding of the effectiveness of coalition loyalty programs. From a practitioner perspective, this paper provides a modeling framework that can easily identify the value of having a store in a coalition network by considering not only the direct impact of participation but also how their participation can influence revenues at other stores as well.

In other words, participating in the network can have a "rippling" effect throughout the network. As will be shown later, an equally important part of deciding which stores to include in the network is the concentration of the network, or how closely the network stores are located relative to each other. As discussed in the counterfactual section, maintaining the centrality of the network is critical when selecting network partners.

A coalition loyalty program can influence cardmember behavior in only a handful of ways. First, the rewards structure can be designed to be attractive enough to motivate card usage. However, once a earnings schedule is in place it may be very difficult or costly to change, and these changes must apply to all cardmembers regardless of individual differences. Another way coalition LPs may be able influence behavior is through marketing actions, assuming that the coalition has at least some control over offer generation and offer aggregation. For instance, the network partners may generate offers for the coalition to disperse to card members. In this setup the coalition has control over which offers go to which customer to maximize response. However, what is more likely is that coalition members will be marketed to at the whim of the coalition partners, and the coalition acts as nothing more than a middleman between the coalition and its members. Of interest in this paper is the last lever that the coalition has control over – the design of the network itself. This decision is least likely to be influenced by the other partners in the network (unless, of course, one partner threatens to leave the network if another joins) and is one way in which the coalition may have direct control over individual preferences. For instance, in a high income neighborhood there may be an emphasis on obtaining a certain mix of stores that might be different from the mix of stores in a low income neighborhood. In the most general form, there is no reason that the coalition can't cherry pick stores that are most likely to lead to revenue maximization.

As previously mentioned, determining the value of a partner is not as straightforward as simply comparing card revenue before and after a store joins the network. The attractiveness of the network to a member should be evaluated more holistically. In other words, the sum may be more that its parts. When entering a coalition network, a store can influence cardmember spend in many ways. The most desirable case is that a store's entry into the network causes the cardmember to infuse additional revenues into the network through that store. However, another potential and far less desirable situation is that the entry of one store causes members to simply direct revenue from another similar store to the store that just joined. In this scenario, the store entry accomplished nothing. In the worst case scenario, a network entrant may actually cause revenue to decline. If the store presents an image that is too counter to the image the cardmember is trying to associate with the cardmember may reduce spend within the network. For instance, if a coalition of high end luxury retailers decides to expand the network to include low end retailers the gains from the smaller retailers may not outweigh the losses incurred from high end shoppers deciding that the coalition is no longer aligned with their interests. For these reasons it is critical that the coalition carefully evaluate both the same store and cross store effects throughout the network.

Understanding the cross-store effects within the coalition network is similar to what has often been described as "earn" and "burn" partnerships (Blattberg et al., 2008). In this relationship, a customer "earns" rewards at one location only to "burn" them at another. One example of this type of program most similar to the concept of the coalition structure studied here was the AOL AAdvantage Reward Program (Direct Marketing, 2000). In this partnership, AOL and American Airlines were earn and burn partners for each other Regan (2000). AOL customers earned airline miles while shopping at their affiliates or by flying on American Airlines. These miles could be spent on either merchandise or free trips. This program was launched in 2000 but was modified substantially in January 2002 by removing the key feature of earning miles while shopping. Even though the success of coalition programs remains mixed partnering has become a common feature of many rewards programs. Many credit cards have programs with multiple partners such as airlines, retailers, and travel agents Polaniecki (2001).

One advantage of studying coalition loyalty programs is that the network can evolve over time. This variation resulting from stores entering and leaving the network can be used to identify the value of participating in the network. This helps to remove some of the self-selection issues that are typically present when analyzing loyalty programs – the card holders do not know ahead of time which stores will enter or leave the network so each entry or exit is essentially an exogenous shock to the program. In a typical loyalty program analysis the analyst only observes from members who have already decided to join that LP. In a coalition LP, each unique network structure can be interpreted as its own LP. For example, consider a hypothetical scenario where there are two stores in the market, Store A and Store B. Suppose there is a coalition network in place that has Store A in the network but excludes Store B. The analyst observes all shopping behavior, both at Store A and at Store B through the cardmember's credit card. If Store B enters the network, the analyst can now compare the cardmember's activity at Store B both before and after it entered the network. In addition, it is easy to see how the LP itself has changed from the cardmember's perspective: at one point they are in a one store LP and now they are in a two store LP. This unique evolutionary behavior of coalition LP's provides an ideal setting to understand the contribution of each store into the network. In addition, the analyst will also be able to determine if total spend changes, or whether revenue simply shifts from A to B or possibly from B to A.

For these reasons understanding the contribution of each store within the network in terms of same store and cross store impact is critical to the success of a coalition loyalty program. The remainder of the paper is organized as follows. First I discuss the coalition dataset used in estimation. Then I introduce the model and discuss the estimation strategy. Next I review the results and illustrate how the model can identify valuable network structures. Finally I conclude.

2.2 Dataset

The dataset comes from a large European coalition program. Customers (also referred to as "cardmembers") sign up for a loyalty program combined with a credit card. The credit card is used for transactions as any other typical credit card, including purchases and cash withdrawal. Purchases can be made (and points accumulated) anywhere credit cards are accepted, however points are earned at a faster rate at "partner" retailers (also referred to as "in-network" stores) compared to "out-of-network" stores. Points are earned according to a schedule and are redeemed for cash vouchers, which can be used to purchase additional goods and services at selected partner retailers. In addition to managing the points tracking and voucher generation and redemption, the firm sends outbound marketing campaigns to encourage sign-up, purchases, and voucher usage. The earning rate of this program is detailed below in Table 10.

In-Network Earning Ratio	1.0 points per local currency unit
Out-of-Network Earning Ratio	.2 points per local currency unit
Points to Voucher Value	500 points = 5 units of local currency
Frequency of Voucher Production	Monthly, if $>$ 500 points
Voucher Units	5, 10, 20, 50, 100
Validity	2 years
Limit of Using Vouchers	100% of purchase amount

Table 10: Coalition Points Schedule

I observe all transactions between from April 2004 through August 2009. Each transaction has details on the amount purchased, the date of the transactions, and how many rewards points were earned. If the purchase was conducted at an in-network store, additional details on the store are given. If the transaction occurred out-of-network, a generic category label is assigned (even if that particular store was previously in the network or will be in the network at a future point in time).

For the model estimation the dataset is grouped into "regimes", or periods in which the network is static. A regime is at most a month long, so if the network remains static for more than one month the data will be split into multiple records. The dataset has 384 regimes over a network of 24 stores and 23 categories. The 24 stores were in the network at some point during the observation period. The dataset is aggregated over 1,403 cardholders, all of whom signed up for the credit card during the data date range. At each regime the key metric of interest is the average daily amount spent on the credit card each store and category. Note that this average is taken across all card members that are currently in the network at the time of that regime, so it is not conditional on cardmembers that have activity on their cards. Table 11 summarizes the dataset.

Date Range	April 2004 through August 2009
Regimes	384
Stores	24
Categories	23
Total Cardholders	1,403
Avg. Daily Spend (across all cardholders)	$8.52

Table 11: Coalition Program Dataset

Figure 14 shows average daily credit card revenue over time across cardholders who were active (i.e., enrolled) at that point in time. I see there there is initially a depressed period of card activity followed by a substantial increase near 2007. This increase coincides with the substantial network expansion as shown in Figure 15.

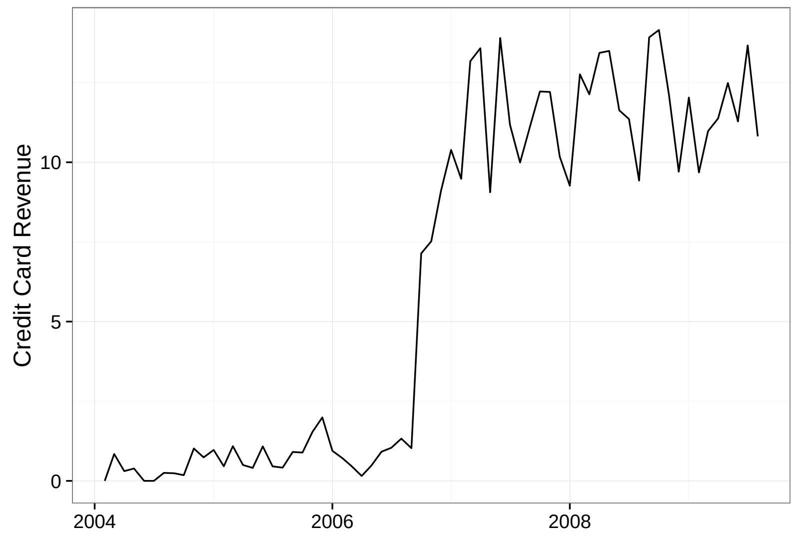


Figure 14: Daily Card Spend (across all card holders)

The evolution of the network in Figure 15 shows that the network is small in 2004 and then expands rapidly in 2007 before again tapering off in the mid-2008. The variation introduced from both no-growth, slow-growth, and fast-growth periods will assist with identifying the impact from individual stores.

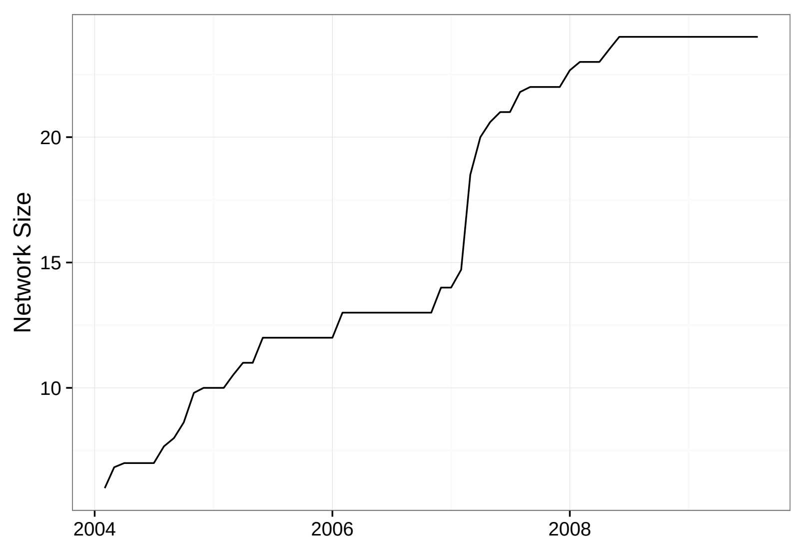


Figure 15: Network Evolution

2.3 Model Free and Preliminary Evidence

The contribution of this research is in providing a model that quantifies the value of an individual store within a coalition loyalty program network. As argued earlier, changing the network structure is the easiest and most direct way a typical coalition program can influence the behavior of its members, so it is critical that managers are able to evaluate the contribution of each store. In this section I provide evidence to suggest that changes in network structure do in fact influence member shopping behavior. This evidence will motivate the need for model development.

Figure 16 illustrates the relationship between average monthly spend from each active cardmember and the size of the network. For a relatively small network, each cardmember generates about $100 in revenue each month on the coalition credit card. As the network grows in size there is a clear increase in the average amount spent per cardmember: with 15 stores in the network each cardmember spends about $200 and at its largest size of 24 stores in the observed data the spend is about $300 per person. There appears to be a clear correlation between the number of stores in the network and the amount spent by each cardmember.

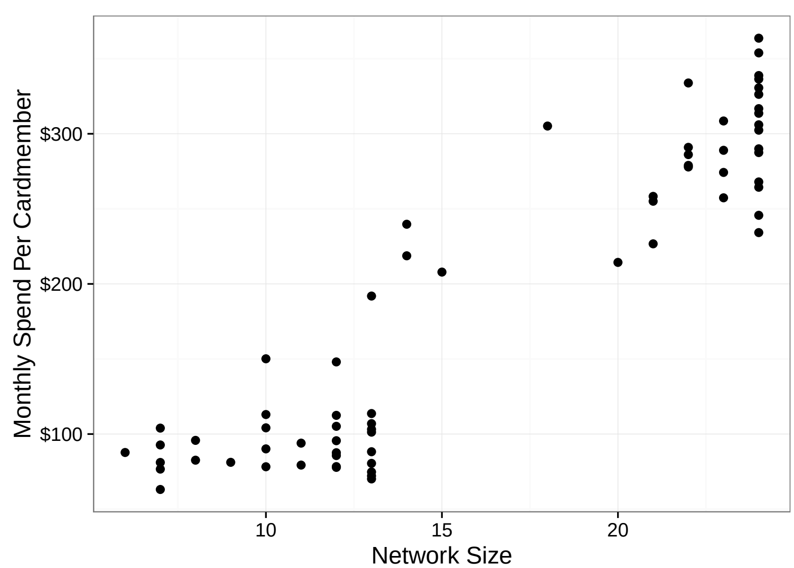


Figure 16: Spend by Network Size

To explore that relationship a bit further, Figure 17 decomposes the average monthly spend above based on whether the spend occurred inside the network (i.e., at one of the stores currently in the network) or outside the network. Interestingly, the in-network spend remains relatively constant even as the network grows but it is the out-of-network stores that benefit most with the change in the network size. As will be shown later, this highlights the importance of accounting for the complete revenue impact of adding stores to the network: even if adding a focal store to the network does not result in increases in spend at that store the network *overall* may benefit.

There are a few potential explanations for this counter-intuitive finding. The first is that perhaps the cardmembers incorrectly infer stores are in the network (or will soon be in the network) when in fact they are not. For example, a customer may notice that more stores are being added to the network and thus simply assume that the stores they tend to shop at will be added as well. This will cause them to shift their spending from other credit cards to the coalition credit card in order to earn higher points as soon as possible. Unfortunately, since activity from *other* credit cards is unavailable this is difficult to test.

Another potential explanation is that the stores that are being added to the coalition increase the brand image of the coalition itself. So even though cardmembers fully understand that they are earning rewards at a lower rate at out-of-network stores the fact that certain stores are being included in the network sends a signal about the strength of the coalition brand. As previously discussed, this would be an example of one type of "ripple" effect that the model can account for: the addition of a particular store may do very little to the revenue at that store but overall card spend may be positively affected through the cross-over effects.

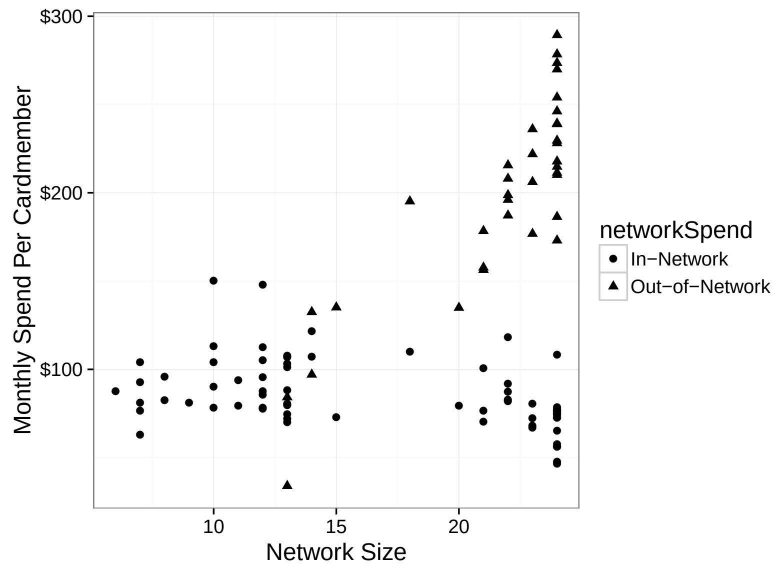


Figure 17: Spend by Network Size and Revenue Source

Finally, another reasonable explanation is that this is simply a time trend and that the growth in the network corresponds to changes in card spend over time. As previously shown, even though the growth in the card spend from Figure 14 appears to correspond with the change in the network size from Figure 15 it is worth examining the time trend more thoroughly.

To determine the predictive strength of the time trend versus the size of the network I regress the log of the average monthly spend per cardmember on the network size and a time trend variable. The results are shown in Table 12 and suggest that it is not simply the mere passage of time that is accounting for these changes in individual level spend but rather the changes in the size of the network itself – as the network grows cardmembers tend to spend more on the card.

Coefficient	**Estimate**	**SE**	**p-value**
Intercept	3.681	0.124	$<$.001*
Network Size	0.088	0.019	$<$.001*
Time	-0.001	0.006	.888
R-Squared	**0.823**		

Table 12: Spend Regression

The evidence presented suggests that the network size itself appears to be driving changes in cardmember behavior. In the next section I introduce the model that will quantify the contribution from each store while accounting for both same store and cross-store effects.

2.4 Model Outline

I model purchase volume using a multivariate normal framework. In each regime the purchase volume of each store and category is a function (among other things) of store and category level coefficients and the cross-effects among stores and among stores and categories.

Let S be the number of stores that have ever been in the network and C the number of store categories designated by the coalition program. Each network store belongs to one of the categories, but not every category needs to be represented by a store. I let $v = S + C$ represent the total number of network stores plus the categories.

For a given individual i at regime time t the purchase volume across stores and categories is represented by:

$$y_{it} = \begin{bmatrix} y_{it}^{1} \\ \vdots \\ y_{it}^{S} \\ y_{it}^{1} \\ \vdots \\ y_{it}^{C} \end{bmatrix}$$

That is, the amount spent in a specific regime t for individual i is represented by a vector of spend at all stores that have ever been in the network at any point plus all category level spend that is *not* associated with a specific store that happens to be in-network at that time.

The managerial problem describe earlier in this chapter calls for us to model the inter-relationships among purchase volumes across the various stores. Given that each store has as an essential characteristic its spatial location on a two-dimensional map, that the effects of purchasing in one store on purchasing in another store may be mediated by the accessibility of one store from another store and finally that spatial co-location is an important driver of accessibility, a natural model to apply here is the spatial autoregressive model, developed in the spatial statistics literature. In the Gaussian conditional dependence form of the spatial autoregressive model, the realization of the response for any one unit is written as a linear function of the realizations of the responses of the other units, with a Gaussian error term added. Optionally, the Gaussian error terms for the various units may be correlated. Denoting the vector of realizations across all units at time t as $\boldsymbol{y_t}$, the spatial autoregressive model can be expressed via the following equation:

$$\boldsymbol{y_t} = r\,\boldsymbol{W}\boldsymbol{y_t} + X_t\beta + u_t$$

The W term corresponds to a square matrix that represents the structure of the spatial contagion, is assumed to be known a priori and is not estimated. The r is a scalar representing the average strength of spatial contagion. The β term is a matrix representing the effect of each covariate X on each component of $\boldsymbol{y}_t$.The values of the r term and the β term are estimated from the data. Note that without the first term in the above expression, conditioning only on the covariates X_t and with the components of the error vector being correlated, we have the Seemingly Unrelated Regressions (SUR) model of Zellner. Therefore, the spatial autoregressive model can be viewed as a generalization of the SUR model.

The equation above gives the conditional form of the spatial autoregressive model, in that each unit's realization is expressed in a form conditional on the realizations of the other units, as can be seen by the fact that the vector y_t appears on both sides of the equation. One can express this model alternatively in the unconditional form, where each unit's realization is written as a function of only the covariates $X_t$and the parameters of the model, but not as a function of the other units' realizations. This is the form we use in the exposition that follows. An important point about the usual spatial autoregressive model is that the contagion structure matrix W is taken to be sparse and known and that the spillover effect r is constant across all pairs of directly spatially connected units. If the contagion structure is highly localized and known, then this approach is defensible. In our problem context, however, we do not have well-formed expectations on the contagion structure. We cannot, for instance, confidently posit that two stores will have high mutual spillover because they are located next to each other or that stores further apart will have lower spillover. In our situation, across-unit covariation will likely be dependent not just on geographic coordinates but also on public transportation connectivity among stores, on complementarity in stores' offerings and on commonalities in shopper segments. For this reason, we take the contagion matrix W to be fully dense, thereby accommodating a wide range of positive and negative contagion structures in a flexible, data-driven manner. Furthermore, we take the spillover strengths to be different across the different pairs of stores, in contrast with the uniform strength r assumed in the traditional model. Finally, we do not constrain the spillover strengths to be symmetric. This allows our model to accommodate the situation where the spillover from one store to another can be different from the second store to the first. This generalized spatial autoregressive model, with a cross-effects matrix that is flexible, dense and potentially asymmetric is what we use in our empirical work.

The purchase volume for a given individual and regime is specified as follows:

$$\log\left(y_{it}\right) \quad \sim N\left(\mu, \Sigma\right) \text{ where } \mu = \begin{bmatrix} \beta_0 + \Gamma\text{Month}_{it} + \beta_{s1} + X_t A_1 \\ \vdots \\ \beta_0 + \Gamma\text{Month}_{it} + \beta_{sS} + X_t A_S \\ \beta_0 + \Gamma\text{Month}_{it} + \beta_{c1} + X_t E_1 \\ \vdots \\ \beta_0 + \Gamma\text{Month}_{it} + \beta_{cC} + X_t E_C \end{bmatrix}$$

$$\text{and } \Sigma = \begin{bmatrix} \sigma^2_{s1} & & & & & \\ & \ddots & & & & \\ & & \sigma^2_{sS} & & & \\ & & & \sigma^2_{c1} & & \\ & & & & \ddots & \\ & & & & & \sigma^2_{cC} \end{bmatrix}$$

β_0 is an intercept term for baseline spending and is constant across all stores and categories. Γ is a vector of eleven month coefficients that capture seasonality, where ΓMonth= $\gamma_2 \mathbb{I}\left[\text{Month} = 2\right] + \ldots + \gamma_{12} \mathbb{I}\left[\text{Month} = 12\right]$. β_{si} is the store specific intercept for store i andβ_{ci} is the category specific intercept for category i. X_t is a $1 \times S$ binary matrix representing store membership of all stores at regime t. A_s is the sth column of the cross store effects matrix A, which is of dimension $S \times S$. This matrix measures how the membership status of each store influences the purchase behavior at other stores in the network. The columns in this matrix represent *outgoing* spillover effects, that is, how does the shopping behavior at other stores change after a target store joins the network. The rows represent the *incoming* spillover effects, that is, how are revenues diverted from other stores into the target store upon joining the network. The diagonals of this matrix represent the same-store effect of being the network versus being out of the network.

E_c is the cth column of the cross-effects between stores and categories matrix E, which is of dimension $S \times C$. This matrix quantifies how the presence of a store in the network shifts spending towards category level spending rather than other in-network stores. For instance, when a target store is in the network this matrix measures the impact of spend at the category level excluding the stores that have ever been in the network.

2.4.1 Missing Data and Imputation

One challenge presented in this dataset is that once a store leaves the network the analyst can only identify the purchase at the category level, not the store level. For example, suppose a specific grocery store (say Whole Foods) decides to join the coalition network. When Whole Foods is in the network any purchases at Whole Foods will be labeled as such. During the periods when Whole Foods is not in the network (either before joining the network or after leaving the network) the analyst would only be able to identify Whole Foods purchases as "Grocery" purchases. Of course, there are other grocery stores for a cardholder to shop at so when Whole Foods is out of the network "Grocery" purchases may be from Whole Foods or they may from another grocery store.

To handle this missing data problem, an imputation process is embedded in the estimation procedure. For a given i and t, y is dimension $v \times 1$ and Σ is a diagonal matrix of dimension $v \times v$. Suppose that I have linear constraints on the y in the form of $Q \times y = h$, where Q is a matrix of size $q \times v$. So h is a column vector of size $q \times 1$. In this application the number of constraints equals the number of stores that are in the network at regime t plus the number of categories. For each store that is currently in the network, the corresponding row in Q will be all zeros except that there will be a one in the position for that store. For each category, each corresponding row of Q will be zeros except that there will be a one for every store belonging to that category and currently *out* of the network, and also a one corresponding to that category's position. The corresponding elements in h will be the expenditure totals for the in-network stores and also the category totals excluding in-network stores.

Given this, the expected value of y given the constraints is:

$$\mu + \Sigma Q' \left((Q\Sigma Q')^{-1} \right) (h - Q\mu)$$

2.4.2 Conditional E-M Algorithm

The imputation process is nested into a conditional EM algorithm, which proceeds as follows:

- *Step 0: Initialization (optional):* First, initialize y_{it} as follows. For every store in the network, set it to be the observed expenditure. For every store not in the network, impute it to be the expenditure averaged over all periods that it *was* in the network. For each category, set the y to be the sum of all expenses in that category in stores outside the network *minus* the amounts imputed in the out-of-network stores in that category.

- *Step 1: Maximization Step:* Given the imputed y_{it} estimate all the model parameters. This is broken into two steps and results in μ and the residual variance leads to Σ.

  - *Step M1:* All parameters *excluding* the crossover effects are estimated: Γ, $\beta_{s1} \ldots \beta_{sS}$, and $\beta_{c1} \ldots \beta_{cC}$. This is done by using a simple ANOVA estimation methodology. After all of the parameters are initialized first get an estimate of the month effects Γ by subtracting out the the intercept, store, and category effects and aggregate by month. Next, subtract out the intercept and newly estimated month effects to get the store and category effects β. Finally, subtract our the newly estimated month, store, and category effects to estimate the intercept. This process is repeated and converges within only a few iterations.
  - *Step M2:* All crossover effects A and E are estimated. First, using the estimated coefficients from Step M1 subtract out the portion of Y attributed to

the non-crossover effects: $Y_{M2} = Y - Y_{M1}$. The joined matrix $[A, E]$ is then estimated by $\left[\hat{A}, \hat{E}\right] = X^{+}Y_{M2}$, where X^{+} is the Moore-Penrose inverse.

- *Step 2: Expectation Step:* Given the estimates of both μ and Σ, the imputation formula above is used to fill in the missing values of y_{it}. The imputed values of y_{it} along with the observed y_{it} gives the complete y_{it}.

The algorithm alternates between Step 1 and Step 2 until convergence. The algorithm is guaranteed to yield consistent estimates of the model parameters. It is important to note, however, that the estimates produced by this procedure do not correspond to maximum likelihood estimates and therefore these estimates suffer some loss of statistical efficiency.

2.5 Results

The estimation algorithm converges relatively quickly, as shown in Figure 18. The tolerance for convergence was an absolute difference of log-likelihoods of less than 1e-10.

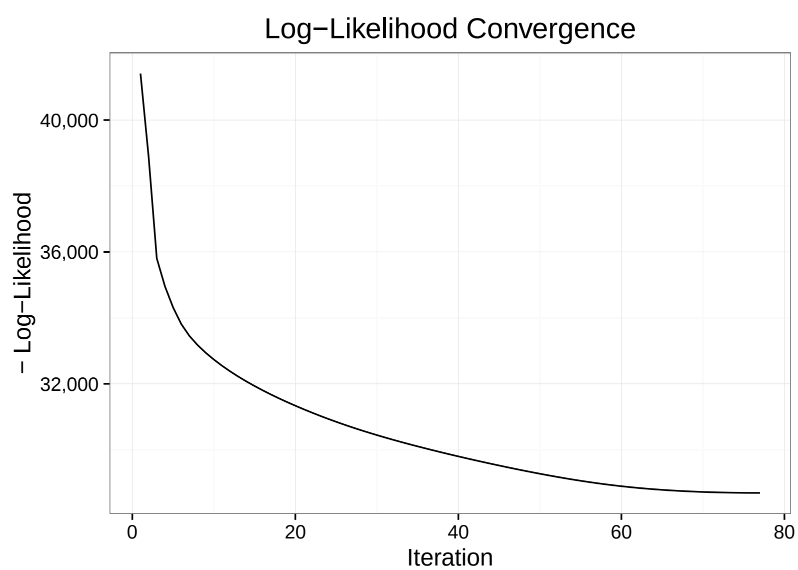


Figure 18: Convergence

Table 13 shows the coefficient estimates for the intercept and month coefficients. For this coalition program, purchase activity tends to be strongest in the Spring and Fall months while slightly depressed in the summer months.

Coef.	Description	Est.
β_0	Intercept	-6.649
γ_2	February	0.339
γ_3	March	-0.092
γ_4	April	0.370
γ_5	May	-0.003
γ_6	June	0.110
γ_7	July	-0.003
γ_8	August	-0.017
γ_9	September	0.310
γ_{10}	October	0.737
γ_{11}	November	0.159
γ_{12}	December	0.134

Table 13: Intercept and Month Effects

Table 14 shows the category level coefficient estimates. Department stores and women's clothing have the highest category spend. The category level variance exhibits a lot of variation across categories. For instance, the variance in the "Clothing" category tends to be much smaller than the variance in the "Airlines" and "Auto Dealer" category, which is intuitive – purchase amounts for clothes are probably much more consistent than those for cars or airline tickets.

Category	β_c	σ_c^2
Airlines	-7.46	16.99
ATM	-6.98	6.89
Auto Dealer	-9.24	20.87
Auto Services	-9.96	13.58
Automobile Rental Agency	-8.25	18.68
Beauty and Barber Shops	-7.09	7.98
Clothing	-7.58	2.86
Cruise	-9.31	12.21
Department Stores	3.51	8.28
Electronics	-7.58	15.09
Entertainment	-8.67	16.64
General Merchandise	-5.98	2.95
Grocery Stores and Supermarkets	-7.31	4.02
Home Goods	-5.29	8.79
Hotels Resorts	-6.67	4.05
Medical	-8.22	18.84
Men's Clothing Accessories	-8.43	18.72
Other	-6.71	6.47
Railways	-7.83	13.40
Restaurant	-6.55	2.43
Service Station	-6.85	3.69
Travel Agencies and Tour Operators	-7.94	18.33
Women's Clothing Accessories	41.01	25.37

Table 14: Category Estimates

Table 15 shows a summary of the 24 store level coefficient estimates. For identification purposes β_1 is set to zero. There is substantial variation in the store impacts, even from stores within the same category. For example, some clothing stores have positive impacts while others have negative impacts. In addition, the variation σ_s^2 fluctuates significantly across the stores. As will be explored later it is not enough to explore these coefficients on their own – the spillover effects need to be accounted for too. That is, even if a store has a negative effect on purchase behavior there may be advantages to keeping a store either in or out of the network depending on its effect on the purchase activity at other stores.

Branch	Category	Join Date	Exit Date	β_s	σ_s^2
1	Women's Clothing Accessories	8/28/2008	-	0.00	1.97
2	Department Stores	2/5/2009	-	1.59	3.57
3	Clothing	3/17/1999	-	7.79	20.34
4	Clothing	1/17/2000	-	-4.31	2.88
5	Clothing	1/18/2000	-	24.09	15.36
6	Clothing	1/18/2000	-	15.04	13.57
7	Clothing	1/19/2000	-	0.16	4.90
8	Clothing	1/19/2000	9/30/2008	8.58	0.56
9	Clothing	1/19/2000	-	42.69	28.91
10	Clothing	3/31/2004	-	-6.22	4.67
11	Clothing	8/4/2004	-	-4.95	3.20
12	Clothing	2/2/2005	-	6.69	8.68
13	Clothing	11/10/2005	-	-3.16	30.60
14	Clothing	12/23/2005	-	7.11	1.23
15	Other	9/28/2005	-	-5.89	0.27
16	Other	11/15/2005	-	-11.82	0.40
17	Clothing	11/18/2005	-	8.80	23.57
18	Women's Clothing Accessories	1/24/2006	-	-6.03	7.36
19	Women's Clothing Accessories	4/6/2006	-	0.19	5.79
20	Clothing	5/4/2006	-	-7.74	0.24
21	Clothing	1/8/2007	-	-3.50	7.15
22	Department Stores	3/19/2007	-	-2.56	3.03
23	Department Stores	2/5/2009	-	-4.92	2.49
24	Women's Clothing Accessories	2/22/2007	-	14.20	10.99

Table 15: Store Level Coefficient Summary

To better understand the store level coefficient estimates, the relationship between β_s and σ_s^2 is plotted in Figure 19. In general, a higher store level effect is associated with a higher variance but there are also a few stores that have a low (or sometimes negative) effect that also exhibit high variance.

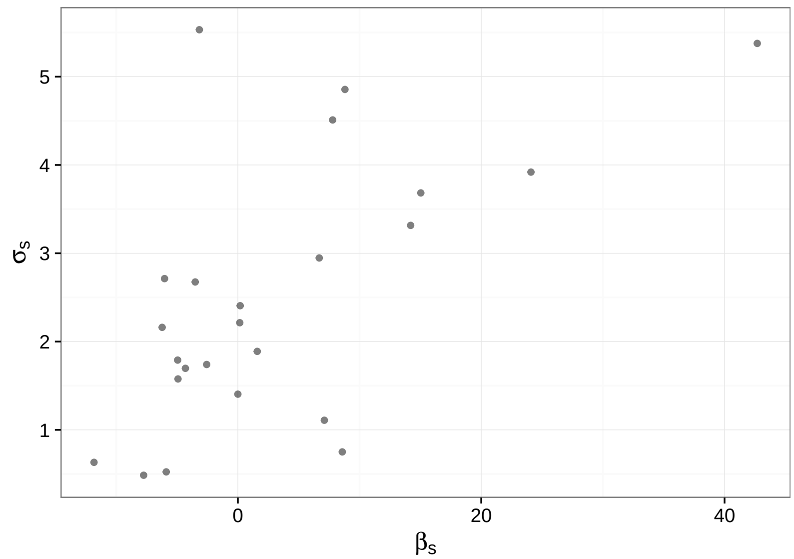


Figure 19: Store Coefficients

Since displaying all of the cross effects from matrices A and E is quite cumbersome the next section works presents a few counterfactuals which illustrates how these cross effects can influence decision making.

2.6 Counterfactuals

Of primary interest from the model estimation are the store-to-store and store-to-category effects from the A and E matrices. These cross effects provide insight as to which stores have the greatest impact on the network by accounting for each store's own effects along with each store's spillover effects across the network. In this section I present a few simple analyses that illustrate how the model can be used to analyze the benefit of the coalition network.

2.6.1 Store Spillover

When a store enters the coalition network it may generate spillover into other stores. Likewise, there may be spillover into the target store from the other stores currently in the network. It is important to recognize that these effects are not necessarily symmetric between stores. For instance, when store A joins the network it may benefit store B at the expense of less revenue for store A. The cross store effects estimated in the A matrix inform the analyst of these asymmetries. Recall in the model specification A_s represents column s from matrix A and when multiplied with the current network status at time t (X_t) the result is an estimate of the average daily revenue over all network participants. Therefore the sth column of A contains the spillover effects *from* or *out of* store s. Likewise, the sth *row* of A contains the spillover effects *into* store s, which represents how much store s benefits from other stores in the network. To designate these differences, I denote ν_s^O as the spillover benefit *out* of store s and ν_s^I as the spillover benefit *into* store s. More formally:

$$\nu_s^O = \sum_j a_{js} \text{ for } j \neq s$$

$$\nu_s^I = \sum_k a_{sk} \text{ for } k \neq s$$

Where a_{ij} is the $[i, j]$th element of matrix A. Then the total *network* level benefit from store s is:

$$\nu_s^T = \nu_s^O + \nu_s^I$$

It is important to note that there is also category level spillover as well, captured in the E matrix. That is, when a store joins the network card members may change spending behavior within specific categories. For instance, if a department store joins the network and is conveniently located near a service station there may be spillover into that category on the card. The category level benefits κ from store s are denoted by:

$$\kappa_s = \sum_k e_{sk}$$

Where e_{ij} is the $[i, j]$th element of E. Given this, the coalition benefit of an individual store can be represented by the spillover effects from the store, plus the spillover effects into the store, plus the category level effects from the store: $\nu^O + \nu^I + \kappa$. Figure 20 shows the distribution of estimated coalition benefits across stores. Some stores have strong negative impacts that reduce the estimated weekly spend to essentially zero while others have strong positive effects.

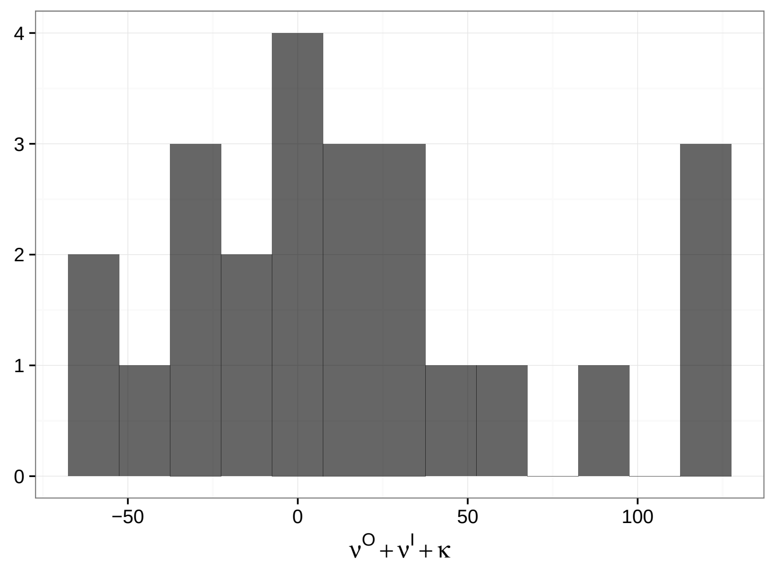


Figure 20: Distribution of Estimated Coalition Benefit

These spillover effects can then be regressed on store level characteristics to determine if there is any pattern associated with the coalition benefit. I regress the coalition benefit on the distance to the city center and the merchant category. I also include a polynomial on the measure of the centrality of each branch, which is the sum of the distance between the target store and all other stores in the network. Thus a lower value means that it is easier to get to other stores within the network from the target store. The results from this regression are shown in Table 16. Interestingly, the coalition benefit is stronger (marginally) as stores move outside of the city center. Also, centrality does play a significant role: the impact from a store is stronger as it becomes closer to other stores. This is intuitive: if a store that joins the network is located very close to a store that is already in the coalition it is easier for card members to take advantage of the a variety of coalition benefits. Thus the centrality of the network appears to play an important role in the effectiveness of the coalition network.

Coefficient	**Est.**	**SE**	**p-value**
Intercept	-1.06	18.19	0.95
log(km to city center)	45.57	25.70	0.09
Category: Department Stores	-27.05	34.23	0.44
Category: Other	12.65	40.67	0.76
Category: Women's Clothing Accessories	-19.57	30.32	0.53
Centrality	-370.45	163.60	0.04
Centrality^2	179.91	107.83	0.11
Adjusted R-Squared	**.07**		

Table 16: Coalition Benefit Regression

2.6.2 Same-Store Effects

Just as a store entering the network can influence the performance of other stores in the network, participation in the network can influence a store's own revenues as well. At first pass it is not immediately clear that joining the network would result in greater revenue from the current card member base. On the one hand, a higher earnings rate should be more attractive at bringing customers to the store. On the other hand, it is possible that the inclusion of certain stores may dilute the brand image associated with the coalition and shift revenues to other other stores that are not currently in the network.

The same-store effects are simply the diagonal elements of the matrix A. Figure 21 shows that the distribution of the same-store effects is generally negative. However, it is important to note that for most stores the *overall* impact is still positive due to the positive spillover both to and from each store.

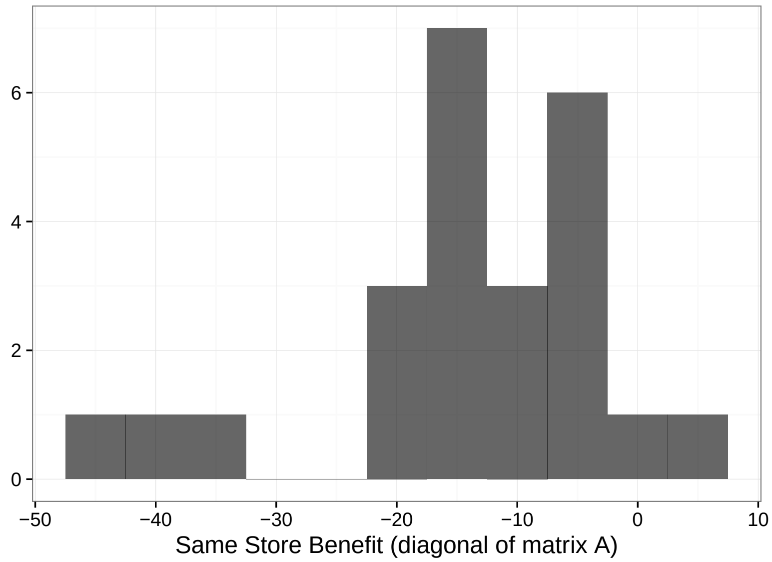


Figure 21: Distribution of Same-Store Benefit

As with before, these store level effects can be regressed against the store level characteristics to determine if there is any pattern in the findings. These findings are shown in Table 17. Unlike the coalition benefit from above (which appears more to be a function of relative location) the same-store effects appear to be solely a function of the store category. Department stores and the "Other" stores show the strongest positive effect. Interestingly, if a manager were to use only the same-store effects and not consider the potential network effects the recommendations on which stores should join the network can be vastly different.

Coefficient	**Est.**	**SE**	**p-value**
Intercept	-17.49	3.59	0.00
log(km to city center)	-3.02	5.07	0.56
Category: Department Stores	14.90	6.76	0.04
Category: Other	19.52	8.03	0.03
Category: Women's Clothing Accessories	9.92	5.99	0.12
Centrality	37.29	32.3	0.26
Centrality^2	6.71	21.29	0.76
Adjusted R-Squared	**.24**		

Table 17: Same Store Regression

2.6.3 Network Constraints

The cross effects can also be useful in situations where there are constraints placed on the network structure. For instance, one reasonable constraint imposed by managers might be that there should only be one store represented from each category. Another potential constraint is that only a certain number of stores can enter the network, thereby placing a cap on the total potential coalition size. For these two types of constraints it is important to understand how the addition or removal of a particular store can have a ripple effect throughout the network. Substantial spillover effects both to and from a target store may outweigh any lackluster same-store effects.

Starting with the first constraint, let's assume that a manager wants to only have one store from each category in the coalition network. In the estimation, the 24 stores are distributed across the following four categories:

Category	**# of Stores**
Clothing	15
Women's Clothing Accessories	4
Department Stores	3
Other	2

This means there are $15 \times 4 \times 3 \times 2 = 360$ potential network combinations where only one store is represented from each category. At this small of scale, it is easy enough to simulate through all network combinations to find the optimal structure under this constraint. To do this the model estimates are used to simulate 1,000 regimes under each potential network structure (where one and only one store is represented from each category). The estimated daily revenue per card member from each regime is averaged together. The result is an estimate of each card member's projected daily spend for each of the 360 constrained network structures. Figure 22 shows the sorted estimated daily revenue across all 360 constrained network structures. The estimated daily spend (across *all* card members) in the best structure is \$5.41 versus \$4.12 for the worst structure. This difference of \$1.29 per day translates to over half a million dollars in annual card revenue for the 1,403 card members in the estimation dataset.

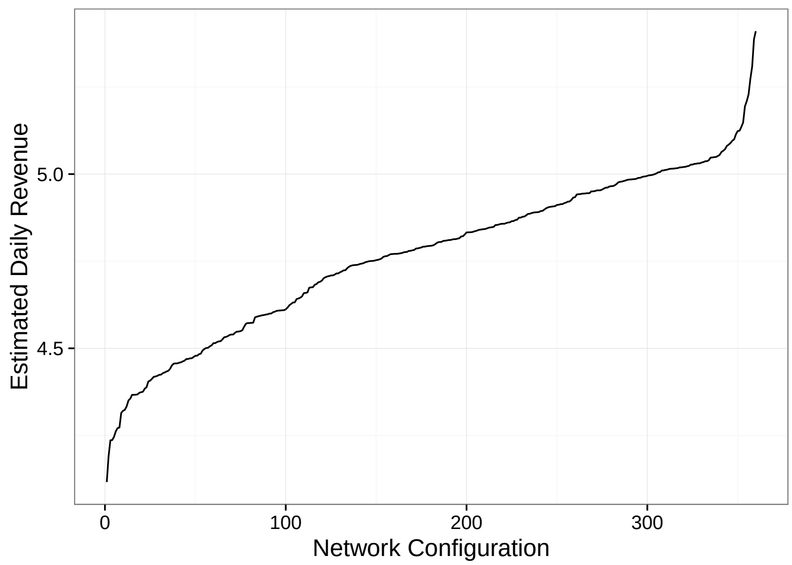


Figure 22: “One per Category” Constrained Network Optimization

The simulation also illustrates the importance of centrality when designing the network. Figure 23 plots the latitude and longitude of the best and worst network structures in this constrained search. In order for the data to remain anonymous the latitude and longitude values have been differenced to zero. Notice that in the best scenario (under the “one per category” constraint) all of the stores are located close to each other while in the worst case scenario the stores are positioned relatively far apart. It is also important to note that the cluster of “best” stores are located in the city center, suggesting that in addition to centrality of the network the location itself matters.

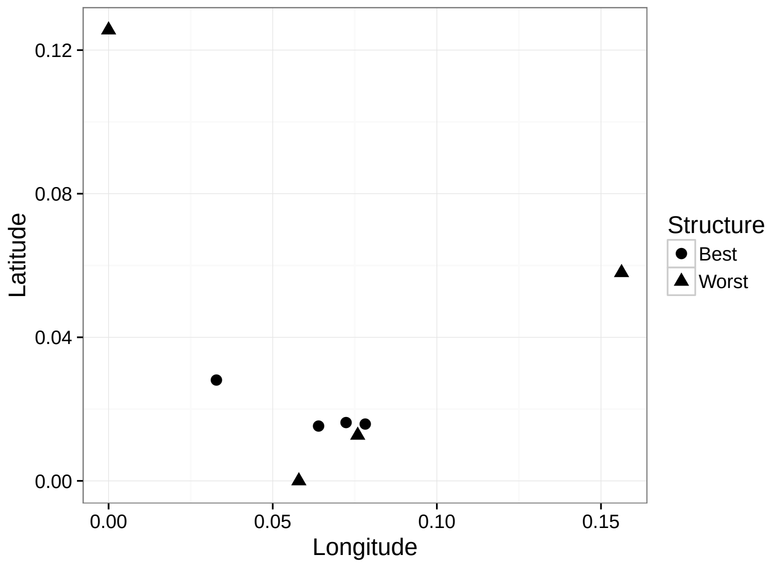


Figure 23: "One per Category" Best and Worst Locations

Accounting for the second constraint adds some complexity to the simulation. In order to find the best network where the size is capped, all possible combinations of stores need to be considered. Suppose there are S possible stores to be in the network, this means there are 2^S store combinations to consider since each store can either be in the network or out. However, with an imposed cap on the network size $\eta < S$ then there are only $\binom{S}{\eta}$ simulations to conduct. Likewise, if the search is over the space with an optimal size $S^* \leq \eta$ then $\sum_{k=0}^{\eta} \binom{S}{k}$ simulations need to occur.

Suppose the goal is to find the most profitable network that consists of up to five stores. In this example, $\eta = 5$, $S = 24$, and $\sum_{k=0}^{\eta} \binom{S}{k} = 55,455$. Figure 24 shows the sorted estimated daily revenue across all potential network structures. The estimated daily spend is \$5.46 for the best structure versus \$3.58 for the worst structure. This difference of \$1.88 per day translates to nearly one million dollars in annual card revenue for the 1,403 card members in the estimation dataset.

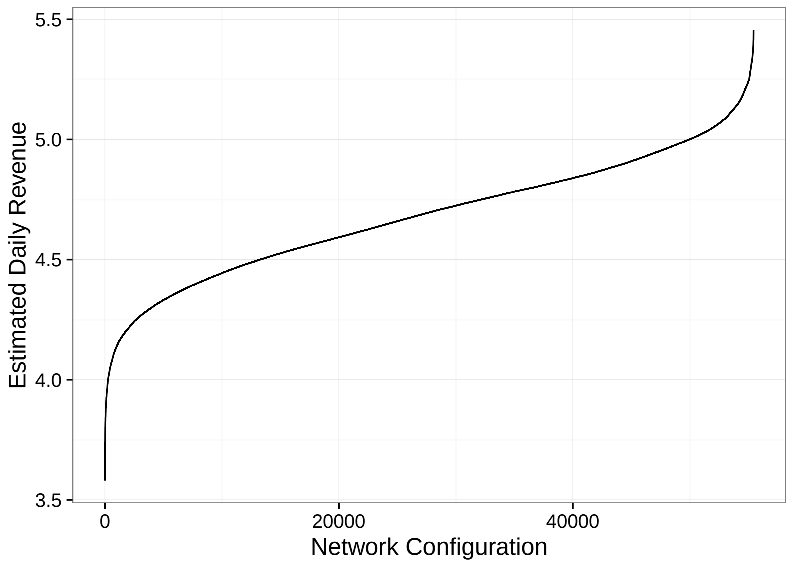


Figure 24: "Up to Five" Constrained Network Optimization

Figure 25 plots the latitude and longitude of the best and worst network structures in this constrained search where the manager wants to consider all possible network configurations of up to five stores. In this simulation, both the best and worst performing strategies included five stores. Notice that as in the last constraint in the best scenario all of the stores are located close to each other while in the worst case scenario the stores are positioned relatively far apart. Even though being located in the city center appears to help, notice that even in the worst network structure there is a store located in the city center as well. This might indicate that looking at location in isolation might be unwise – it is the relativity of the locations in the network that appears to matter.

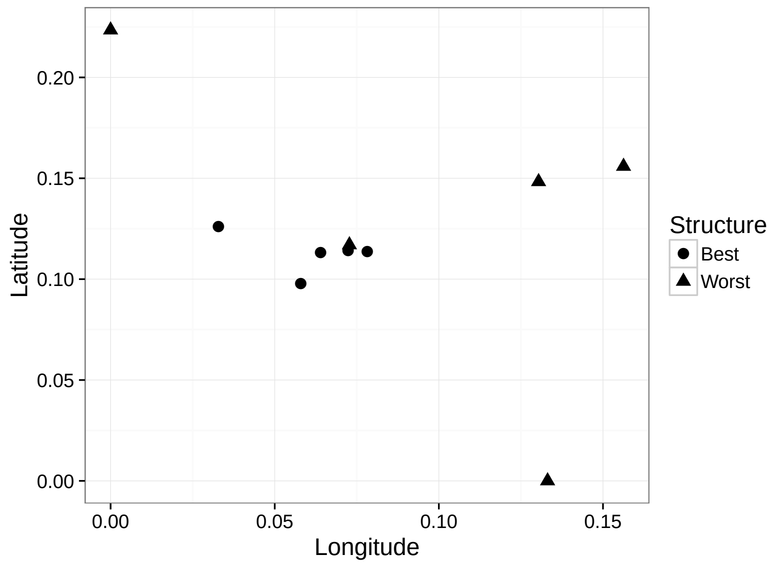


Figure 25: "Up to Five" Best and Worst Locations

2.7 Conclusion

In this paper I presented a model that quantifies the benefit of stores entering a coalition network. This research provides further insight into understanding the effectiveness of coalition loyalty programs and in particular identifying the value of each store within the network. An important insight for managers is that the value of a store entering a network depends on both 1) how revenue will change at the focal store and 2) how that presence of the focal store will influence spending to and from other stores currently in the network. This research shows that these spillover effects can be substantial and sometimes even greater in magnitude that the same-store effects. The model can assist mangers in determining which stores should be considered when modifications are to be done on the network structure.

This research is not without its limitations. One challenge is that for large coalition networks estimation becomes intractable and understanding any effects beyond first order interactions is difficult.

3 Extension to Essay 1 and Essay 2: Departures from Bayesian Rationality

3.1 Introduction

The way in which humans form beliefs, make judgments, and decide on a course of action has long been a source of interest for researchers from many fields including psychology, management, economics, and computer science, to name a few. The quest to understanding and replicating human thought itself has endured over many decades as the implications of this could have dramatic repercussions. For marketers, understanding how humans respond to past experiences is key to developing better targeted marketing strategies.

Bayesian inference has remained central in the debate of understanding how exactly the brain operates. The appeal of this method is not hard to argue: it is simple, intuitive, and applicable to a wide range of problems. At its core, Bayesian inference is a method of updating beliefs as new evidence presents itself. These beliefs are represented though probability distributions. This is contrast to frequentist inference, which examines how frequently one might expect to observe an outcome by emphasizing proportionality in the data.

Not surprisingly, there is a strong ongoing debate as to whether or not the brain truly operates in a Bayesian optimal way or rather if some other mechanism is at play. In this paper I explore some of these potential departures from rationality. While Bayesian inference has been found to be a good approximation to human behavior in many applications, many argue that this representation how of the brain operates is not evidence that the brain is in fact processing data in a Bayesian fashion.

This extension builds upon the first two essays. In the first essay, I assumed that gamblers learn about slot machine payoffs through Bayesian updating. As will be discussed, there are other learning models that may more appropriately represent how gamblers integrate new information with their past gambling experience. In addition, in the second essay the implicit assumption is that customers are fully aware of the network structure at all times such that changes in the network are immediately recognized. However, it is not unreasonable to argue that there is a timing lag in how customers learn about the network structure which may explain at least a portion of the observed behavior. This extension is meant to more deeply explore alternative learning models and more specifically the process of integrating new information with prior knowledge. A more thorough understanding of the learning process may influence the implications of the models for managers.

The paper is organized as follows. First I provide a review of Bayes' theorem. Then I review prior literature and highlight the conflicting points of view. Next I summarize the potential departures from rational Bayesian inference. After this, I propose potential modeling solutions that will capture these departures from rationality and discuss some of the estimation challenges. Building on the proposed models I discuss the ideal data structures that could estimate these models. Finally I conclude.

3.2 Bayes' Theorem

In this section I provide a quick review of Bayes' theorem and discuss some of its advantages and limitations. Bayes' theorem calculates the probability of an event conditional on other events that may occur.

$$P(B|E) = \frac{P(E|B)\,P(B)}{P(E)}$$

In words, the probability of event B occurring given that event E is true is a function of the probability of event E occurring given that event B is true along with the probabilities of observing E and B without regard to each other. With respect to belief updating, the theorem can be interpreted as finding the probability that a "Belief" is true given "Evidence", where $P(B)$ and $P(E)$ represent the prior probabilities of both the belief and evidence, respectively. In other words, the prior belief is "updated" as new evidence comes forward – the result being the posterior belief $P(B|E)$. This posterior belief then takes the place of the prior until even more evidence is obtained. The more evidence is assessed the better the judgments become.

It is difficult to argue against the intuition of the theorem: humans start with beliefs that evolve as more evidence is realized. That being said, there are a few well-known examples where the results from Bayes' theorem diverges from most people's expectations (the following examples from Bain (2016)). For example, many people suggest that when flipping a fair coin getting all heads is or all tails is less likely that getting something, say, tails-tails-heads-tails-heads. It is not and Bayes' theorem explains why: since the coin tosses are independent no one sequence is more likely than another. There is also the Monty Hall problem. Participants are asked to pick one of three doors (A, B, or C) and behind one of them is a prize. When the host opens one of the non-winning doors (say C) the contestant has the opportunity to stick with their original door (say A) or switch to the other unopened door (say B). Most people will think that switching doors will not make a difference but if Bayes theorem is applied switching doors actually improves your chance of winning. Finally, Bayes' theorem reveals that even a test that is 99% accurate may be wrong half the time when telling people about a rare condition. This counter-intuitive explanation can be explained by the fact that the probability of the condition is so low. Say 10,000 people are tested for a rare disease, and only 1% (100 people) have the condition and the other 9,900 do not. Of the 100 people that have the disease, a 99% accurate test will detect 99 of the true cases, leaving a single false negative. But a 99% accurate test also produces false positives at a rate of 1%, meaning that of the 9,900 people that do not have the condition, 1% of them (99 people) will be incorrectly informed that they have the disease when in fact they do not. That means that of the 198 people who were informed they have a rare disease, only 99 of them actually do. Thus the probability that a positive test results from is a true positive is only 50%.

Even in light of these (and other) counter-intuitive findings many researchers remain adamant in their argument that people's brains process information in a Bayesian fashion. In the next section these arguments are explored in more depth before proposing alternatives that depart from Bayesian rationality.

3.3 Literature Review

Researchers have long argued over whether or not Bayesian inference is an accurate representation of the human mind. While there appears to be consensus that Bayesian inference is a reasonable *representation* of how the brain works, it is unclear whether it is truly accurate of how humans process information. This section reviews some of the main arguments both for and against Bayesian inference as a representation of the human mind. It is not meant to be a comprehensive review but rather a summary to highlight the primary arguments on either side of the debate. At a very general level, the division in opinion seems to pivot on how calculating the brain truly is as it integrates new information with past experiences. Generally, those who support Bayesian inference tend to assume that there is some (usually fixed) truth that must be learned over time. Because of this, past experiences can matter quite a bit. This is problematic for those who do not support Bayesian inference because this implies that there is some "state" that is being kept track of which is updated over time. Reinforcement learning does not necessarily have a notion of "states" and is better suited to allow for changes in reality itself. In other words, rather than try to find the true state of the world humans simply integrate positive and negative experiences in specific ways.

Payzan-LeNestour and Bossaerts (2011) argue that the main advantage of the Bayesian learning framework is that unlike other learning frameworks (such as reinforcement learning) it explicitly tracks three notion of uncertainty. This is because Bayesians form a model of the world that differentiates between risk (outcomes are uncertain, even with perfect knowledge), estimation uncertainty (the payoff probabilities themselves need to be learned), and unexpected uncertainty (the payoff probabilities may change over time). Under reinforcement learning, only the value of a chosen option is updated on the basis of the reward (or loss) prediction error – the difference between the received and the anticipated reward (or loss) (Rescorla, 1972). No attempt is made to disentangle the different sources of prediction error and usually the learning rate is kept constant. The authors do not attempt to augment the reinforcement learning model with an ambiguity penalty because they point out that non-Bayesians do not sense ambiguity. They argue that since the representation of ambiguity is absent in the context of model-free reinforcement learning, ambiguity cannot weigh in the exploration strategy. For this reason one should not combine model-free reinforcement learning with an ambiguity penalty/bonus. However, they did concede that full Bayesian updating is reflected in human learning only if enough structural information of the outcome generating process is provided. Specifically, the ability to track unexpected uncertainty, and hence, to detect jumps in the outcome probabilities, appeared to rely on instructions that such jumps would occur. When participants were not informed about the presence of unexpected uncertainty, their choices could equally well be explained in terms of simple reinforcement learning.

According to Li et al. (2017) a general theme in memory research is that humans have limited cognitive resources, and must make decisions – either consciously or unconsciously – about what information to maintain, and what to discard. They note that previous research in resource rationality suggests that individuals make optimization decisions according to a utility function, where predictive power is maximized and cost is minimized. While this describes the behavior of an ideal Bayesian agent, it fails to capture the behavior of a standard learner. Instead, recent research shows that individuals do not make predictions based on the full posterior probability distribution, but rather base decision on samples from this distribution.

Bayesian-belief updating involves two elements: a prior, which represents belief states before observing data, and a likelihood function, which links observed evidence with beliefs by assigning probabilities. Cassey et al. (2016) quantitatively examined the degree to which human behavior approximated Bayesian inference at the level of the individual subject, using the well-studied paradigm of simple probabilistic inference. When these investigations have been applied at the level of individuals, the comparison between human and Bayesian inference has been largely qualitative (see Williams and Griffiths (2013)); that is, analysis questions are frequently of the type, "Do the participants' responses move in the direction predicted by Bayesian inference?" Their experimental paradigm reduced the inference task to a level that allowed the Bayesian model of cognition to be quantitatively compared with individual participants' behavior. They manipulated the difficulty of a simple prediction about an alien who was flipping coins. Participants were asked about the nature of a coin (i.e., fair or biased) both before and after seeing a sequence of outcomes. The authors recruited a sufficiently large number of participants that the full ranges of prior and posterior beliefs were sampled, and also enough that they were able to analyze important subsets of the data. The authors found that participants' inferences were mostly inconsistent with Bayes' rule. It is well-documented that humans have a propensity to discount the value of initial information in favor of novel information. Their analysis highlights the importance of testing cognitive theories at a quantitative level. The Bayesian theory of cognition, which is apparently successful when tested at a qualitative level, fails when tested quantitatively.

Bowers and Davis (2012) argue that there is no evidence that people are in fact Bayesian. They first show that the empirical evidence for Bayesian theories in psychology is weak, then show that the empirical evidence for Bayesian theories in neuroscience is weaker still, and finally they challenge the general scientific approach that characterizes Bayesian theorizing in cognitive science.

Generally, it is often assumed that evolution produces systems that satisfice (Simon,

1956) or meliorize (Dawkins, 1982). That is, selective adaptation produces "good enough" solutions, or "better than alternative" solutions, but not optimal solutions. The Bayesian approach, by contrast, appears to claim that evolution has endowed us with brains that are exquisitely good at learning and exploiting the statistics of the environment, such that performance is close to optimal. Bowers and Davis argue that the flexibility of Bayesian models, coupled with the common failure to contrast Bayesian and non-Bayesian accounts of performance, has led to a collection of Bayesian "just so" theories in psychology and neuroscience: sophisticated statistical analyses that can be used to explain almost any behavior as (near) optimal. If the data had turned out otherwise, a different Bayesian theory would have been carried out to justify the same conclusion, that is, that the mind and brain support near-optimal performance. In other words, simply because a Bayesian learning model can fit the data does not imply that the underlying process that occurs in the brain takes the same form.

In addition, they argue that the necessity of considering prior probabilities (base rates) is one that most people find quite counter-intuitive. Of course, the idea that previous knowledge influences perception is not novel. For example, there are abundant examples of top-down influences on perception. What is unique about the Bayesian hypothesis is the claim about how prior knowledge is combined with evidence from the world, that is, the claim that humans combine these sources of information in an optimal (or near-optimal) way, following Bayes's rule. Again, the output of a good heuristic model might approximate the decisions of a Bayesian model in a given context, but it will not approximate the underlying processes. As noted by Sloman and Fernbach (2008), Bayesian models developed in this manner should be considered descriptive, not rational.

They continue to argue that in order to provide some evidence in support of the claim that the mind relies on Bayesian-like algorithms, it is necessary to show that these algorithms do a better job than non-Bayesian models in accounting for human performance. However, this is rarely done. Consider the phenomenon of probability matching. When asked to make predictions about uncertain events, the probability with which observers predict a given alternative typically matches the probability of that event occurring. For instance, in a card guessing game in which the letter A is written on 70% of the cards in a deck and the letter B is written on 30%, people tend to predict A on 70% of trials (e.g., Stanovich (1999)). This behavior is irrational, in the sense that it does not maximize utility. The optimal (Bayesian) strategy would be to predict the letter A every time, as this leads to an expected outcome of 70% correct, whereas perfect probability matching would predict an accuracy rate of only 58%.

One point that they try to emphasize is that advocates of Bayesian theories often show strong confirmation bias: Theorists take the successful predictions of a Bayesian model as support for their approach and ignore the fact that alternative non-Bayesian theories might account for the data just as well, and sometimes better. The consequence of this is that Bayesian models tend to receive credit for relatively trivial predictions, which could be derived from any theory that assumes performance is adaptive, or else are modified with post hoc assumptions when the data do not follow what would be predicted on the basis of the rational analysis alone.

There is relatively little evidence in support of methodological and theoretical Bayesian theories in psychology and neuroscience: Bayesian theories are so flexible that they can account for almost any pattern of result; Bayesian models are rarely compared to alternative (and simpler) non-Bayesian models, and when these approaches are compared, non-Bayesian models often do as good a job (or better) in accounting for the data; and the authors argue that the neuroscience data is ambiguous at best.

Still, Bowers and Davis note that it is important to recognize some key contributions of the Bayesian literature. First, Bayesian modelers have highlighted to the importance of characterizing the environment when developing theories of mind and brain. Second, Bayesian theorists have highlighted the importance of addressing why questions, the authors do take the point that too many non-Bayesian theories in cognitive science can be characterized as an exercise in curve fitting. Third, and related to the second point above, Bayesian modelers have emphasized top-down or function-first strategies for theorizing, such that theories are first constrained by the complex computational challenges that individuals face. By contrast, many, if not most, non-Bayesian models are first constrained by a limited set of data in a given domain, with the goal of elaborating a theory to accommodate more complex phenomena after capturing the more simple phenomena first.

Through a series of experiments Mozer et al. (2008) also find evidence that people are not Bayesian. They argue that the findings supporting Bayesian inference set forth by Griffiths and Tenenbaum (2006) reflect the averages over many individuals. By testing many different heuristic algorithms they obtain fits that are as good as, or sometimes better than, the Bayesian model implies. In one algorithm they find that individuals treat the task as a memory retrieval task, and base their response on similarity of the query to the stored instance. Further testing finds that other models outperform Bayesian inference in other domains. They argue that a preference for one model or the other cannot be justified on grounds of data fit alone. Rather than finding support that an inidividual's mind are Bayesian and utilize prior distributions to draw complex inferences their results are consistent with the another less "dramatic" possibility: that our minds reason from only a small number of instances (less than three) and that the mechanisms of reasoning may be simple heuristic algorithms. They concede by agreeing that it is often very illuminating to view human reasoning from a Bayesian perspective but an overemphasis on the ways in which reasoning conforms to Bayesian principles may draw attention away from important psychological distinctions and may obscure important memory and processing limitations of human reasoning.

Charness and Levin (2005) also present evidence that is inconsistent with Bayesian learning, noting that people may often ignore prior information when forming beliefs. Based on their experiments, they advocate reinforcements learning, where one is more likely to pick actions associated with successful past outcomes. Interestingly, they find that when payoff reinforcement and Bayesian updating are aligned, nearly all people respond as expected; on the other hand, when these forces clash and people are paid for their initial choice, nearly half of all decisions are inconsistent with Bayesian updating.

Interestingly, there is some recent support that individuals are Bayesian from a biological perspective. Kolossa et al. (2015) find empirical support of the Bayesian brain hypothesis. They find that the brain appears to code and compute the distinguishable aspects of Bayes-optimal probabalistic inference. They do not claim that the mathematics behind Bayesian inference is consciously available to the participants but suggest instead that probability theory is essentially common sense reduced to mathematics.

Finally, many researchers analyze human performance on bandit problems. The bandit problem is a dynamic decision-making task that is simply described, well-suited to controlled laboratory study, and representative of a broad class of real-world problems. In bandit problems, people must choose between a set of alternatives, each with different unknown reward rates, to maximize the total reward they receive over a fixed number of trials. A key feature of the task is that it challenges people to balance the exploration of unfamiliar choices.

Steyvers et al. (2009) point out that Bandit problems have a known optimal solution for which human decision-making can be compared and how well people solve optimization problems like bandit problems may provide a window onto differences in basic human cognitive ability. By studying this class of problems, they are able to observe individual differences which make it useful to be able to measure the extent to which people adhere to optimal decision processes, rather than simpler heuristics. The authors showed that the Bayesian methods worked well, even with a small amount of data

Lee et al. (2011) also studies bandit problems and tests a variety of reinforcement learning algorithms such as the "Win-Stay Lose-Shift" heuristic, the "epsilon-greedy" heuristic from reinforcement learning where the decision-maker controls the balance between exploration and exploitation, "epsilon-decreasing" (a variant on the epsilon-greedy) where the probability of exploration decreases as the trials progress, and the "π-first" model which assumes two distinct stages in decision-making. In the first stage, choices are made randomly. In the second stage, the alternative with the greatest estimated reward rate is always chosen. The first stage can be conceived as "exploration" and the second stage as "exploitation".

In addition they also develop a new model of bandit problem decision-making, motivated by the idea of latent states used by the π-first model. The development comes in two parts: first they implement and evaluate a "full" latent state model, that allows for switching between exploration and exploitation at any trial in a bandit problem. They show, however, by applying this model to the human and optimal decision data, that it is possible to simplify the model significantly. Accordingly, they finish this section defining a simple latent state model that can subsequently be compared to the simple benchmark models already described.

They find that most participants, in most conditions, begin with complete exploration, and transition at a single trial to complete exploitation, which they maintain for all of the subsequent trials. They note that this general finding is remarkable, given the completely unconstrained nature of the model in terms of exploration and exploitation states. All possible sequences of these states over trials are given equal prior probability, and all could be inferred if the decision data warranted. They suggest that three basic challenges in studying any real-world decision-making problem are to characterize how people solve the problem, characterize the optimal approach to solving the problem, and then characterize the relationship between the human and optimal approach. Their results show how the use of simple heuristic models, using psychologically interpretable decision processes, and based on psychologically interpretable parameters, can aid in all three of these challenges.

3.4 Departures from Rationality

The primary contribution of this paper is to evaluate the potential departures from the rational Bayesian inference framework and propose alternatives that better capture these departures. More concretely, these departures influence how the posterior evolves. In this section I discuss potential departures from Bayesian inference and in the next section discuss more specifically how to model these departures.

3.4.1 Posterior Recall

In Bayesian inference, it is assumed that that individuals have perfect recall of their posterior. Depending on the frequency, strength, or recency of the signals they experience it is not unreasonable to suggest that individuals may forget their updated posterior at the time the next signal is experienced. For example, if a gambler is learning about slot machine payout rates then posterior recall is likely to be very high between spins during a gaming session. However, the recall between the last spin on one trip to the casino and the first spin on the return trip to that casino many months later is likely to not be as strong. Whether individuals use the true posterior or their *recalled* posterior can have substantial implications on the proposed learning process under certain situations.

3.4.2 Measurement Error in the Data

In Bayesian inference the updated posterior becomes the next period's prior which is combined with a new signal (i.e., data) to generate the following updated posterior. In Bayesian inference it is assumed that the signal is recorded by the human without error. However, there may be situations where the data is interpreted or processed in a way consistent with measurement error.

In some cases, the magnitude of the error may be a function of the prior itself. For example, suppose that an individual has a very precise and very high prior on a signal.

If a relatively low signal arrives the individual may simply discount this data and fail to update their posterior in a typical Bayesian fashion. The implication is that the updated posterior is consistent with a different signal rather than the true signal. That is:

$$
\begin{aligned}
p(B|E^*) &\propto p(E^*|B)\,p(B) \\
E^* &= E + \varepsilon
\end{aligned}
$$

The evidence used in the updating mechanism is actually E^*, not the true signal E. In this example the error is a function of the prior $p(B)$ but the construction of the error term could take many other forms. Rather than be a function of the prior, it may be a function of the signal itself (e.g., extreme signals may cause more error), an individual difference (ε_i), random noise, or a function of any other characteristics surrounding the delivery of the signal.

Clearly, an error in the updated posterior will simply flow through to an error in the next period's prior. So the measurement error has a "double" effect on the posterior – affecting the interpretation of the current period's data through $p(E|B)$ and also potentially influencing how the prior $p(B)$ has evolved from signals in the prior periods.

As an example of measurement error, consider the "antennaegate" issue Apple faced in 2010.[9] When Apple released their iPhone 4 it became apparent very quickly that the phone's reception was significantly impaired when a user covered the exposed antennae band. The unusual placement of the antennae was justified by Apple as a way to keep the case slim. Then CEO Steve Jobs responded to the crisis by stating that users should simply not hold the phone in such a way to alter the reception.

Translating this event into the rational Bayesian inference perspective, the posterior belief on Apple's quality should combine the new data (the malfunctioning antennae) with

[9] http://www.pcworld.com/article/201297/apples_iphone_4_antenna_gate_timeline.html

consumers' prior beliefs (say, that Apple produces high quality products) such that the updated belief on quality should be lowered. However, both the extremity of the event and the power of Apple's brand may affect how the data becomes integrated with the priors. For simplicity, assume that the strength (or precision) or the prior beliefs across consumers is constant and the only difference is in the location. That is, some consumers think Apple produces high quality products while others believe the products are low quality. For consumers with the high quality prior, the antennae issue may simply be discounted or ignored because this is such an unusual event or they choose to ignore it because of their love for Apple. The implication is that the signal is interpreted as a far less negative event than it truly was and for these Apple fans both the location and precision of the posterior may remain unchanged. But consumers with the low quality prior may use this as further evidence that Apple does not produce quality products – the effect being a reduction in the posterior quality and an increase in the precision of the posterior. In other words, the consumers with low quality priors are now even *more* certain that Apple produces *even worse* products than they previously believed. This is an example of how the measurement error is a function of both brand loyalty and the prior belief in quality.

3.4.3 Multiplicative versus Simpler Mechanisms

The multiplicative property of Bayesian inference causes "later" signals to influence the beliefs relatively less and less. However, there may be simpler mechanisms such as reinforcement learning (discussed in further detail later) that more accurately reflect the underlying updating process. For instance, perhaps the posterior evolves in an additive way. Or rather than evaluating the arithmetic mean the updating process is better reflected through geometric or harmonic means. It is interesting to note that exploring these alternative specifications highlights the potential arbitrariness of the current structure of Bayesian inference (besides, of course, the mathematical convenience of doing so).

However, while it is important to explore the fit of these alternative specifications it cannot be ignored that these modifications to the updating process do not necessarily get us closer to understanding how the brain actually operates, which is one of the most pressing arguments against Bayesian inference in the first place. Given this, alternatives in model specification may need to be augmented with theory in order for the proposed alternatives to hold merit.

3.4.4 Multiplicative but not in Proportion to Fisher Information

In Bayes' theorem the posterior is proportional to the prior times the likelihood. Taking the second derivative of both sides yields that the *information* of the posterior is proportional to the *information* in the prior times the *information* in the likelihood. However, under some modifications to the likelihood this relationship changes. For instance, if the likelihood contains a learning parameter α such that rather than a likelihood of $p(y|\theta)$ I have $p(y|\theta)^{\alpha}$ the resulting information is not the same. As will be shown later, the introduction of a learning parameter in this way requires that the joint posterior of the mean and variance is modeled, rather than the mean conditional on the variance.

3.4.5 Weighting of Observations

In Bayesian inference, the strength of each new signal remains constant. In other words, the only variability between signals is in the value of the signal itself. However, it is not unreasonable to suggest that some circumstances will lead to a different weighting of the signal by the individual. For example, a signal that is unreasonably large or small may be discounted by the individual as outliers. Or the circumstances surrounding the environment in which the signal is generated may lead to biases as well. Returning to the slot machine payout example, a specified payout rate may be given more or less consideration depending on the magnitude of the bet placed. Signals generated from small bets may be weighted in a different way than signals generated from larger bets.

Clearly, for any fair bet the size of the wager should not affect the odds. However, the size of an individual's wager may inadvertently affect their updating mechanism.

As alluded to earlier, the weighting could also be a function of recency of prior signals – a signal that is experienced immediately between two signals (in terms of timing) may receive less attention then a single stand alone signal experienced after a long absence of signals. In traditional Bayesian inference the timing of signals does not influence the updating process but I argue that there may be a situations where it is reasonable to expect this to be true.

To be clear, the weighting of the signals is different from the discussion earlier of the measurement error. Here, the individuals do are not interpreting the true data incorrectly. Rather they recognize that the data for what it is but choose to place larger or smaller emphasis on the signal for a variety of reasons.

3.4.6 Reinforcement Learning

From a behavioral psychology perspective, reinforcement is simply a consequence of a desired outcome. After experiencing a desired outcome, the individual behaves in a way that maximizes these reinforcing consequences and minimizes punishing consequences. This has been applied in machine learning under the name "reinforcement learning". The spirit of both is similar: agents take actions that maximize some type of cumulative reward. As discussed earlier, this is a departure from Bayesian inference in a number of ways. Most generally is that there is no "state" of truth that a human is trying to learn – they simply take actions that reinforce any positive behavior. This allows for the reality to evolve over time rather fixing the truth at a given point for humans to hone in on over time. As will be outlined later in the model development, one advantage of this approach over Bayesian inference is that the order of outcomes matters. In Bayesian inference all past signals are aggregated into the prior belief in such a way that order of outcomes is irrelevant. However, in reinforcement learning the result of a, say, negative-positive signal sequence can be different from that of a positive-negative signal sequence.

3.5 Modeling Departures from Rationality

In this section I briefly outline the modeling techniques that can account for some of these departures from rationality. I focus on the following three departures: 1) how measurement error may affect the interpretation of the signal, 2) the introduction of learning rates in the likelihood, and 3) reinforcement learning.

3.5.1 Measurement Error in the Data

As previously discussed, an individual may interpret signals with measurement error. This error can be a function of properties of the signal itself, the prior beliefs, or characteristics outside of the updating process itself (such as brand loyalty). The implication of this error is that not only is the current period's posterior affected but all future updating is affected as the posterior becomes the next period's prior. In this way the measurement errors can compound over time.

Broadly, the consequence of errors of measurement is a failure to identify the parameter of interest. Attempting to fix the error can be very complex. For most econometric applications, one option is to simply remove the variable and risk an omitted variable bias. Clearly that is not an option here because the signal itself is key to modeling the updating process. Another option is to form additional assumptions about the measurement error or to obtain additional information that may provide insight into the severity of the measurement error.

Generally, in regards to Bayesian inference rather that update the posterior with the signal E the individual instead updates the posterior with a function of the signal:

$$E^* = f(E)$$

This function captures two components of measurement error in the signal: random errors and systematic errors. Systematic errors are not determined by chance and result in a nonzero error component, which cannot be eliminated by averaging over many observations. A systematic error is simply an offset of the true value – identifying the degree of this offset would assist the analyst in correcting any biases introduced through systematic errors. Random errors are more difficult to account for because as the name suggests the measured value cannot be predicted in such a way that an allowance for the effect could be made.

Allowing δ to represent the systematic error and ε represent a zero mean random error results in

$$f\left(E\right) = \delta + \varepsilon$$

The challenge is in identifying the systemic error δ. As discussed earlier, the systematic error is not necessarily constant and may be a function of either the underlying signal, individual characteristics, or the prior belief. Depending on the quality of the data in the analysis, it may be relatively simple to determine if the systematic errors are a function of other observables in the data. By comparing the posterior distributions across signals and individuals one can identify how the characteristics influence the updating process and either 1) correct for these biases or 2) develop policies based on these biases.

The much more difficult task arises when the systematic error is not linearly related to observed characteristics. Part of the difficulty relates to one of the larger criticisms against Bayesian inference in that with enough parameters an updating process can be fit to any dataset. Of course, it cannot be forgotten that this is also true of most models employed by researchers whether it be a regression or logit. One suggestion is to then draw upon research from other fields such as psychology in order to explain any unusual deviations in the measurement errors. For example, for casino gamblers if the incoming signal is positive or negative outcomes loss aversion informs us that losses loom larger than gains

(Kahneman and Tversky, 1979). If a large gain in model fit results by incorporating this behavioral structure into the model it may be reasonable to suggest that the measurement error is simply a reflection of an underlying psychological process. Even with these options, obviously the most desirable solutions are either better data or field experiments that could accurately identify measurement errors and clearly associate them with an underlying mechanism.

3.5.2 Incorporating Learning Rates

In this section I outline how learning rates can be incorporated into learning models. Suppose that there is a learning rate α that affects the strength in which the likelihood is integrated with the prior when updating beliefs. In this specification, I use the prototypical example of estimating the mean of a population from a sample where y is a vector of n observations from a univariate normal distribution $N(\mu, \sigma^2)$. The likelihood is given by the usual:

$$\begin{aligned} p\left(y|\mu, \sigma^2\right) &\propto \sigma^{-n} \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i-\mu)^2\right) \\ &= \sigma^{-n} \exp\left(-\frac{1}{2\sigma^2}\left[(n-1)\,s^2 + n\,(\bar{y}-\mu)^2\right]\right) \end{aligned}$$

where

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(y_i-\bar{y})^2$$

is the sample variance of the y_i's. The sufficient statistics are $\bar{y}$ and s^2. The conjugate prior density must have the product form $p(\sigma^2)\,p(\mu|\sigma^2)$, where the marginal distribution of σ^2 is scaled inverse-χ^2 and the conditional distribution of μ given σ^2 is normal (Gelman et al., 2004). A convenient parameterization is given by the following specification:

$$\mu|\sigma^2 \sim N\left(\mu_0, \sigma^2/\kappa_0\right)$$
$$\sigma^2 \sim \text{Inv-}\chi^2\left(\nu_0, \sigma_0^2\right)$$

which corresponds to the joint prior density

$$p\left(\mu, \sigma^2\right) \propto \sigma^{-1}\left(\sigma^2\right)^{-(\nu_0/2+1)} \exp\left(-\frac{1}{2\sigma^2}\left[\nu_0\sigma_0^2 + \kappa_0\left(\mu_0 - \mu\right)^2\right]\right)$$

Multiplying the prior joint density by the normal likelihood yields the posterior density

$$\begin{aligned} p\left(\mu, \sigma^2|y\right) &\propto \sigma^{-1}\left(\sigma^2\right)^{-(\nu_0/2+1)} \exp\left(-\frac{1}{2\sigma^2}\left[\nu_0\sigma_0^2 + \kappa_0\left(\mu_0 - \mu\right)^2\right]\right) \\ &\quad \times\sigma^{-n} \exp\left(-\frac{1}{2\sigma^2}\left[(n-1)\,s^2 + n\left(\bar{y} - \mu\right)^2\right]\right) \\ &= \sigma^{-1}\left(\sigma^2\right)^{-(\nu_n/2+1)} \exp\left(-\frac{1}{2\sigma^2}\left[\nu_n\sigma_n^2 + \kappa_n\left(\mu_n - \mu\right)^2\right]\right) \end{aligned}$$

where it can be shown that

$$\begin{aligned} \mu_n &= \frac{\kappa_0}{\kappa_0 + n}\mu_0 + \frac{n}{\kappa_0 + n}\bar{y} \\ \kappa_n &= \kappa_0 + n \\ \nu_n &= \nu_0 + n \\ \nu_n\sigma_n^2 &= \nu_0\sigma_0^2 + (n-1)\,s^2 + \frac{\kappa_0 n}{\kappa_0 + n}\left(\bar{y} - \mu_0\right)^2 \end{aligned}$$

The parameters of the posterior distribution combine the prior information and the information contained in the data. Notice that the posterior mean is a weighted average of the prior mean and the observed mean where the weighting is based on the relative precision of the two pieces of information. The posterior degrees of freedom, ν_n, is the prior degrees of freedom plus the sample size. The posterior sum of squares, $\nu_n \sigma_n^2$, combines the prior sum of squares, the sample sum of squares, and the additional uncertainty conveyed by the difference between the sample mean and the prior mean.

In most applications of Bayesian inference, the analyst estimates the conditional posterior distribution of μ, given σ^2, which is proportional to the joint posterior density above with σ^2 held constant. This leads to the well-recognized formulas for the distribution of the posterior mean.

$$\begin{aligned} p\left(\mu|\sigma^2, y\right) &\sim N\left(\mu_n, \sigma^2/\kappa_n\right) \\ &= N\left(\frac{\frac{\kappa_0}{\sigma^2}\mu_0 + \frac{n}{\sigma^2}\bar{y}}{\frac{\kappa_0}{\sigma^2} + \frac{n}{\sigma^2}}, \frac{1}{\frac{\kappa_0}{\sigma^2} + \frac{n}{\sigma^2}}\right) \end{aligned}$$

Here, if n is large then the posterior distribution is largely determined by σ^2 and the sample mean $\bar{y}$.

Now suppose that there is a learning rate α that affects the strength at which the likelihood is integrated with the prior in forming updated posterior beliefs. That is,

$$p\left(y|\mu, \sigma^2\right) \propto \left[\sigma^{-n} \exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^{n} \left(y_i - \mu\right)^2\right)\right]^{\alpha}$$

The challenge is that when modeling a Bayesian updating process if the analyst is interested in the conditional posterior distribution of μ, given σ^2 then the learning parameter is unidentified along with σ^2. Because of this, the analyst needs to model the *joint* normal probability, not the conditional in order to properly account for the learning rate. The learning rate α is then identified by comparing the changes in the posterior mode from *earlier* observations to changes in the posterior mode in *later* observations.

For a given learning rate α, the usual posterior distribution we saw earlier gets modified to have the following parameter values:

$$\begin{aligned}
\mu_n &= \frac{\kappa_0}{\kappa_0 + n\alpha}\mu_0 + \frac{n\alpha}{\kappa_0 + n\alpha}\bar{y} \\
\kappa_n &= \kappa_0 + n\alpha \\
\nu_n &= \nu_0 + n\alpha \\
\nu_n\sigma_n^2 &= \nu_0\sigma_0^2 + (n\alpha - 1)\, s^2 + \frac{\kappa_0 n\alpha}{\kappa_0 + n\alpha}\,(\bar{y} - \mu_0)^2
\end{aligned}$$

The learning rate mediates the relationship between the prior belief distribution, the observed data and the beliefs of the individual as reflected in that individual's actions. Two individuals with the same observed data but different learning rates will arrive at different beliefs, which will theoretically be reflected in different observed decision behaviors. Therefore, putting together the decision behaviors and the observed data allows us to identify both the learning rate and the parameters of the prior belief distribution. This argument is contingent on the assumption that in the posterior belief distribution, the prior distribution's parameters are separable from the learning rate given distinct enough values of the sufficient statistics $\bar{y}$ and s^2, an assumption that can be verified by inspection of the expressions given above.

3.5.3 Order Effects in the Reinforcement Model

In this section I outline a reinforcement model that accounts for potential ordering effects in the signals, which are not considered in Bayesian inference. In Bayesian inference, the order does not matter at all – each past signal can in essence be exchanged with another past signal. However, for reasons previously mentioned, this might be an unreasonable assumption in some circumstances.

Let p_t be a column vector of length I giving the purchase probabilities for the I items in the choice set at time t. Let δ_i be used to denote a column vector which is zero everywhere except in the ith position where it has a 1. Denote $r(i,a)$ the event that product i receives a reinforcement of strength a. The Suppes model (Suppes and Atkinson, 1960) says:

$$p_{t+1}|r(i,a) = (1-a)\,p_t + a\delta_i$$

If this is followed with a reinforcement of strength b to product j, then I have:

$$p_{t+1}|\left\{r(i,a), r(j,b)\right\} = (1-b)(1-a)\,p_t + (1-b)\,a\delta_i + b\delta_j$$

Obviously this is different from

$$p_{t+1}|\left\{r(j,b), r(i,a)\right\} = (1-a)(1-b)\,p_t + (1-a)\,b\delta_j + a\delta_i$$

Therefore the reinforcement operator is not commutative.

Note that for very small matters of a and b I have $ab \approx 0$, implying that the order doesn't matter:

$$\begin{aligned} p_{t+1}|\left\{r(i,a), r(j,b)\right\} &\approx (1-a-b)\,p_t + a\delta_i + b\delta_j \\ &\approx p_{t+1}|\left\{r(j,b), r(i,a)\right\} \end{aligned}$$

This suggests that the Suppes formulation is revised as follows. For a sequence of operators $r(i,a), r(j,b), r(k,c), \ldots$ then the new probabilities are posited as:

$$p_{t+1}|\left\{r(i,a), r(j,b), r(k,c)\right\} = (1-a-b-c)(1-a)\,p_t + a\delta_i + b\delta_j + c\delta_k$$

Apart from communtativity, the model has the following property: when two reinforcements of strength a and b are applied to the same product i:

$$
\begin{aligned}
p_{t+1} | \{ r(i,a), r(i,b) \} &= (1-a-b)\, p_t + a\delta_i + b\delta_i \\
&= (1-(a+b))\, p_t + (a+b)\, \delta_i
\end{aligned}
$$

This is of course equivalent to a single reinforcement of strength $a+b$. The additivity of reinforcements can be seen as both good and bad: good because of the mathematical simplicity, bad because of the departure from Weber's Law (see Fechner (1966)). The Suppes suggestion of dealing with multiple reinforcements by introducing new levels of reinforcement strength allows one to respect Weber's Law but increases the number of parameters.

3.6 Data

In this section I discuss the properties of data that could identify these departures from rational Bayesian inference. First, the ideal dataset would have a direct reading on the posterior belief. Most often, researchers rely on actions taken *based on* an individual's posterior beliefs. The obvious challenge with doing so is that there is much more noise in the measurement. For example, a typical analysis may try to model purchase behavior as a function of posterior beliefs. This is difficult though because there are many other variables that may influence purchase behavior as well, such as price, promotions, current inventory, or seasonality among others. However, if the analyst has direct access to the individual's posterior beliefs an incredible amount of noise is likely to be eliminated. For example, net promoter scores could be an accurate proxy for beliefs. Ideally, this would be measured after each customer encounter or even at regular intervals. Since many of the departures mentioned focus specifically on the nuances that influence posterior belief evolution having a direct measure of the belief would aid tremendously in estimation. Augmenting the net promoter scores with surveys would provide an even stronger measure of the true belief.

Secondly, a long time series is needed to uncover a) order effects and b) the learning rate α. In order to properly identify order effects, it is necessary for the individual to have exposure to a variety of signals. With only a limited number of signals, the order effects will be difficult to distinguish from the learning process itself. Likewise, the learning rate is only identified through changes in the posterior distribution over time. These changes may be subtle so an extended time horizon can help to identify the learning parameter.

Finally, an accurate reading of the experience itself would fix the measurement error issues previously mentioned. In other words, is there both an objective and subjective measure of the experience? That is, there may be a disconnect between the experience itself and how the individual *perceived* the experience. Having this knowledge would inform the analyst as to how the priors influence the reading of the data by the consumer. For a complete picture, the analyst would need a) the true signal, b) the customer's belief about the signal, and c) the updated posterior of the customer. The comparison between how the posterior *should* be updated and how the posterior is *actually* updated by the individual will provide insight into any biases they are introducing to the updating process as a function of the properties of the signal itself.

3.7 Conclusion

Understanding the way in which individuals make decisions under uncertainty has been of interest to researchers for decades now. At this point, it is unclear if there will ever be one method that shines above the rest. That being said, much attention has centered around Bayesian inference and specifically whether or not it truly reflects the underlying learning process taking place in the human brain. The appeal of this approach due to its mathematical simplicity and intuitiveness is matched against concerns that it is an unrealistic demand placed on the brain and that rather humans rely on even simpler heuristics when processing new information and updating beliefs.

In this paper I highlight the advantages and disadvantages of Bayesian inference and outline potential departures from the rational Bayesian framework. These departures are meant to deal with some of the pitfalls in Bayesian inference in terms of how accurately they can represent the human mind. I also show how to model a few of these departures and discuss the types of data that would assist in identifying these pitfalls.

This paper is not intended to settle the dispute on whether or not humans are Bayesian. Rather, the purpose is to further our understanding of the updating process by proposing alternatives to the traditional specification. In the first essay, the learning mechanism was assumed to be that of traditional Bayesian inference. In the second essay, customers were assumed to have complete knowledge of the changing network structure in real time. However, as this extension has illustrated, these may be very limiting assumptions and the true learning process may take on a very different form. A better understanding of the human decision process and models that can account for potential departures from Bayesian rationality may lead to improvements in a firm's targeting strategy and the design of their loyalty programs.

References

American Gaming Association (2013). 2013 State of the States: The AGA Survey of Casino Entertainment.

Bain, R. (2016). Are our brains bayesian? *Significance*, 13(4):14–19.

Bajari, P., Benkard, C. L., and Levin, J. (2007). Estimating dynamic models of imperfect competition. *Econometrica*, 75(5):1331–1370.

Berman, B. (2006). Developing an effective customer loyalty program. *California Management Review*, 49(1):123–148.

Blattberg, R., Kim, B., and Neslin, S. (2008). *Database Marketing: Analyzing and Managing Customers*. Springer, New York, USA.

Bolton, R. N. (1998). A dynamic model of the duration of the customer's relationship with a continuous service provider: The role of satisfaction. *Marketing science*, 17(1):45–65.

Boulding, W., Kalra, A., Staelin, R., and Zeithaml, V. A. (1993). A dynamic process model of service quality: from expectations to behavioral intentions. *Journal of marketing research*, 30(1):7.

Bowers, J. S. and Davis, C. J. (2012). Bayesian just-so stories in psychology and neuroscience. *Psychological bulletin*, 138(3):389.

Capizzi, M. and Ferguson, R. (2005). Loyalty trends for the twenty-first century. *Journal of Consumer Marketing*, 22(2):72–80.

Cassey, P., Hawkins, G. E., Donkin, C., and Brown, S. D. (2016). Using alien coins to test whether simple inference is bayesian. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 42(3):497.

Charness, G. and Levin, D. (2005). When optimal choices feel wrong: A laboratory study of bayesian updating, complexity, and affect. *The American Economic Review*, 95(4):1300–1309.

Chib, S. and Greenberg, E. (1996). Markov chain monte carlo simulation methods in econometrics. *Econometric theory*, 12(03):409–431.

Ching, A. T., Erdem, T., and Keane, M. P. (2013). Invited paper-learning models: An assessment of progress, challenges, and new developments. *Marketing Science*, 32(6):913–938.

Compton, J. (1999). Players clubs 101.

Dawkins, R. (1982). The extended phenotype: The gene as the unit of selection. san francisco, ca: W h.

DeWulf, K., Odekerken-Schröder, G., and Iacobucci, D. (2001). Investments in consumer relationships: A cross-country and cross-industry exploration. *Journal of Marketing*, 65(4):33–50.

Dorotic, M., Fok, D., Verhoef, P., and Bijmolt, T. (2011). Do vendors benefit from promotions in a multi-vendor loyalty program? *Marketing Letters*, 22(4):341–356.

Eckstein, Z., Horsky, D., and Raban, Y. (1988). An empirical dynamic model of optimal brand choice. *Foerder Institute of Economic Research, Working Paper*, (88).

Erdem, T. and Keane, M. P. (1996). Decision-making under uncertainty: Capturing dynamic brand choice processes in turbulent consumer goods markets. *Marketing Science*, 15(1):1–20.

Fechner, G. (1966). Elements of psychophysics. vol. i.

Ferguson, R. and Hlavinka, K. (2007). The colloquy loyalty marketing census: sizing up the us loyalty marketing industry. *Journal of Consumer Marketing*, 24(5):313–321.

Gelman, A., Carlin, J. B., Stern, H. S., and Rubin, D. B. (2004). *Bayesian Data Analysis*. Chapman and Hall.

Griffiths, T. L. and Tenenbaum, J. B. (2006). Optimal predictions in everyday cognition. *Psychological science*, 17(9):767–773.

Hartmann, W. and Viard, B. (2008). Do frequency reward programs create switching costs? a dynamic structural analysis of demand in a reward program. *Quantitative Marketing and Economics*, 6(2):109–137.

Hotz, V. J., Miller, R. A., Sanders, S., and Smith, J. (1994). A simulation estimator for dynamic models of discrete choice. *The Review of Economic Studies*, 61(2):265–289.

Imai, S., Jain, N., and Ching, A. (2009). Bayesian estimation of dynamic discrete choice models. *Econometrica*, 77(6):1865–1899.

Kahneman, D. and Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica: Journal of the Econometric Society*, 47(2):263–292.

Keane, M. P. and Wolpin, K. I. (1994). The solution and estimation of discrete choice dynamic programming models by simulation and interpolation: Monte carlo evidence. *The Review of Economics and Statistics*, 76(4):648–672.

Kilby, J. and Fox, J. (1998). *Casino Operations*. Wiley & Sons.

Kolossa, A., Kopp, B., and Fingscheidt, T. (2015). A computational analysis of the neural bases of bayesian inference. *Neuroimage*, 106:222–237.

Lee, M. D., Zhang, S., Munro, M., and Steyvers, M. (2011). Psychological models of human and optimal performance in bandit problems. *Cognitive Systems Research*, 12(2):164–174.

Li, L., Supaniratisai, G. P., and Zhang, S. (2017). Resource rational analysis of memory use in bayesian updating.

Li, X. (2014). Product offerings and product line length dynamics. *Available at SSRN 2462577*.

Marketing, D. (2000). Aol and american airlines create largest online customer loyalty program. *Direct Marketing*, 63(1):10.

Meyer, R. J. (1981). A model of multiattribute judgments under attribute uncertainty and informational constraint. *Journal of Marketing Research*, pages 428–441.

Moore, G. and Sekhon, H. (2005). Multi-brand loyalty cards: A good idea. *Journal of Marketing Management*, 21(5-6):625–640.

Mozer, M. C., Pashler, H., and Homaei, H. (2008). Optimal predictions in everyday cognition: The wisdom of individuals or crowds? *Cognitive Science*, 32(7):1133–1147.

Nair, H., Misra, S., Hornbuckle, W. J., Mishra, R., and Acharya, A. (2013). Big data and marketing analytics in gaming: Combining empirical models and field experimentation.

Narayanan, S. and Manchanda, P. (2009). Heterogeneous learning and the targeting of marketing communication for new products. *Marketing Science*, 28(3):424–441.

Narayanan, S. and Manchanda, P. (2012). An empirical analysis of individual level casino gambling behavior. *Quantitative Marketing and Economics*, 10(1):27–62.

Osborne, M. (2011). Consumer learning, switching costs, and heterogeneity: A structural examination. *Quantitative Marketing and Economics*, 9(1):25–70.

Oxford Economics (2014). Economic impact of the us gaming industry.

Park, H. M. and Manchanda, P. (2015). When harry bet with sally: An empirical analysis of multiple peer effects in casino gambling behavior. *Marketing Science*, 34(2):179–194.

Payzan-LeNestour, E. and Bossaerts, P. (2011). Risk, unexpected uncertainty, and estimation uncertainty: Bayesian learning in unstable settings. *PLoS Comput Biol*, 7(1):e1001048.

Polaniecki, R. (2001). Credit cards: How rewarding. *Credit Union Management*, 24(4):30–36.

Regan, K. (2000). Aol trumps amazon with air miles plan.

Reiss, P. C. (2011). Structural workshop paper-descriptive, structural, and experimental empirical methods in marketing research. *Marketing Science*, 30(6):950–964.

Rescorla, R. A. (1972). A theory of pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. *Classical conditioning II: Current research and theory*, pages 64–99.

Roberts, J. H. and Urban, G. L. (1988). Modeling multiattribute utility, risk, and belief dynamics for new consumer durable brand choice. *Management Science*, 34(2):167–185.

Roos, J. M., Mela, C. F., and Shachar, R. (2013). Hyper-media search and consumption. *Available at SSRN 2286000*.

Rossi, P. (2014). *Bayesian Non-and Semi-parametric Methods and Applications*. Princeton University Press.

Rossi, P. E., Allenby, G. M., and McCulloch, R. E. (2005). *Bayesian Statistics and Marketing*. John Wiley & Sons.

Rust, R. T., Inman, J. J., Jia, J., and Zahorik, A. (1999). What you don't know about customer-perceived quality: The role of customer expectation distributions. *Marketing Science*, 18(1):77–92.

Shugan, S. (2005). Brand loyalty programs: Are they shams? *Marketing Science*, 24(2):185–193.

Simon, H. A. (1956). Rational choice and the structure of the environment. *Psychological review*, 63(2):129.

Sloman, S. and Fernbach, P. M. (2008). The value of rational analysis: An assessment of causal reasoning and learning. *The probabilistic mind: Prospects for rational models of cognition*, pages 485–500.

Sriram, S., Chintagunta, P. K., and Manchanda, P. (2015). Service quality variability and termination behavior. *Management Science*, 61(11):2739–2759.

Stanovich, K. E. (1999). *Who is rational?: Studies of individual differences in reasoning.* Psychology Press.

Steyvers, M., Lee, M. D., and Wagenmakers, E.-J. (2009). A bayesian analysis of human decision-making on bandit problems. *Journal of Mathematical Psychology*, 53(3):168–179.

Suppes, P. and Atkinson, R. C. (1960). *Markov learning models for multiperson interactions*, volume 5. Stanford University Press.

Tversky, A. and Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *science*, 185(4157):1124–1131.

Van den Berg, G. J. (2001). Duration models: specification, identification and multiple durations. *Handbook of econometrics*, 5:3381–3460.

Verhoef, P. (2003). Understanding the effect of customer relationship management efforts on customer retention and customer share development. *Journal of Marketing*, 67(4):30–45.

Williams, J. J. and Griffiths, T. L. (2013). Why are people bad at detecting randomness? a statistical argument. *Journal of experimental psychology: learning, memory, and cognition*, 39(5):1473.