Big Data and Marketing Analytics in Gaming: Combining Empirical Models and Field Experimentation*

Harikesh S. Nair
Prof. of Marketing
Stanford GSB

Sanjog Misra
Prof. of Marketing
Chicago Booth School of Business

William J. Hornbuckle IV
President and Chief Marketing Officer
MGM Resorts International

Ranjan Mishra
Senior Partner and President
ESS Analysis

Anand Acharya
Principal, ESS Analysis

Previous versions: Dec 20, 2013; Oct 7, 2014. This version: Oct 26, 2015

Abstract

This paper reports on the development and implementation of a large-scale, marketing analytics framework for improving the segmentation, targeting and optimization of a consumer-facing firm's marketing activities. The framework leverages detailed transaction data of the type increasingly becoming available in such industries. The models are customized to facilitate casino operations and were implemented at the MGM Resorts International's group of companies. The core of the framework consists of empirical models of consumer casino visitation and play behavior and its relationship to targeted marketing effort. Important aspects of the models include incorporation of rich dimensions of heterogeneity in consumer response, accommodation of state-dependence in consumer behavior, and controls for the endogeneity of targeted marketing in inference, all issues that are salient in modern empirical marketing research. As part of the framework, we also develop a new approach that accommodates the endogeneity of targeted marketing. Our strategy is to conduct inference separately across fixed partitions of the score variable that targeting is based on, and may be useful in other behavioral targeting settings. A novel aspect of the paper is an analysis of a randomized trial implemented at the firm involving about 1.5M consumers comparing the performance of the proposed marketing-science based models to the existing status quo. We find the impact of the solution is to produce about \$1M to \$5M incremental profits per campaign, and about an 8% improvement in the Return on Investment of marketing dollars. At current levels of marketing spending, this translates to between \$10M and \$15M in incremental annual profit in this setting. More generally, we believe the results showcase the value of combining large, disaggregate, individual-level datasets with marketing analytics solutions for improving outcomes for firms in real-world settings. We hope our demonstrated improvement from analytics adoption helps accelerate faster diffusion of marketing science into practice.

Keywords: marketing, promotions, casinos, behavioral targeting, nonrandom targeting, endogeneity, field experiments.

*Some numbers in the paper have been scaled or disguised to preserve confidentiality. The views discussed here represent that of the authors and not of Stanford University or the University of Chicago. The usual disclaimer applies. We thank participants at the 2014 Big Data Marketing Analytics Conference at Chicago-Booth, NASSCOM Innotrek in Menlo Park, SICS at Berkeley, Successful Applications of Customer Analytics Conference at WCAI-Wharton, Theory & Practice Conference at Kellogg, Workshop on Social and Business Analytics at McCoombs-UT Austin; at the 2015 AMA Winter Conference in San-Antonio, MSI Regional Knowledge Networking Event in San Francisco, Google-Marketing EDGE I-MIX Program in Mountain View; seminar participants at Columbia, Univ. of Delaware, Univ. of Iowa, Stanford and Wharton; our discussants Vineet Kumar and Puneet Manchanda; as well as J-P. Dube, Pedro Gardete, Chris Nosko, Thomas Otter, Stephan Seiler, Jeff Zweibel and especially Gary Russell, Josh Swissman, and Eric Bradlow and Gunter Hitsch in particular for helpful comments and suggestions. Please contact Nair (`harikesh.nair@stanford.edu`) or Misra (`sanjog.misra@chicagobooth.edu`) for correspondence.

1 Introduction

The advent of "Big Data," and the associated ability to track and measure the behavior of consumers has had a disruptive effect on many industries particularly in the way marketing is conducted and evaluated. By improving the ability to micro-target consumers, and by driving the rise of "evidence-based management" in which decisions are supported by data, the measurability of marketing has improved and several issues like advertising and promotions are now routinely treated as quantitative problems. We describe a marketing analytics system we developed in one industry − gaming and gambling − where transactional-level data on consumer play behavior along with targeted marketing information at fine levels of resolution are now abundant. We show that combining the richness of the data with empirical models of consumer behavior and a state-of-the art optimization system improves the Return on Investment (ROI) of marketing effort at the firm, and increases the profitability of targeted marketing.

In a large-scale randomized field evaluation involving about 1.5M consumers in the firm's database, we find the new system produces between \$1M to \$5M dollars of incremental profits per campaign compared to the status-quo policy. The source of the improvement arises from shifting marketing dollars away from average consumers who would have played even in the absence of the promotion towards marginal consumers for whom the promotion has an incremental impact; and from the improved matching of promotion types to consumer types. Computing an ROI per dollar spent, we find the new policy provides a net ROI of about 2.75 compared to 2.55 for the status-quo approach. Thus, a dollar spent in promotions generates about 20 cents more incremental spending under the new policy compared to the current practice at the firm. Assuming this difference is the best estimate of the incremental profit from the new model, this translates to approximately between \$10M and \$15M in incremental profit from shifting from the status-quo to the new model, assuming the same level of marketing spends. Taken together, we believe these numbers suggest the new policy is successful, and serves to demonstrate the value of marketing analytics in the field.

We developed the new marketing analytics system in collaboration with ESS Analysis, a consulting company, for implementation at MGM Resorts International (henceforth "MGM"), a large gaming and hospitality company based in Las Vegas, NV. The firm manages 11 casinos in Las Vegas including well known portfolio brands like The Bellagio, MGM Grand, Mandalay Bay and The Mirage. The engagement at MGM began in 2010. The models were implemented at MGM in late 2011. The randomized experiments to evaluate the new model were implemented in Spring and Summer of 2012. The project fits into the rising trend of analytics transforming customer facing industries, including the gaming industry. The project is part of MGM's initiatives to use industry-leading analytics to improve the consumer experience at their properties, and to optimize the allocation of the right set of promotions to the right set of their customers.

The core of the framework is built on empirical models of consumer behavior all of which have their genesis in the marketing science literature. While the models are tailored to the casino and gaming setting, important aspects of the models that cut across contexts include incorporation of rich dimensions of heterogeneity in consumer response, accommodation of state-dependence in consumer behavior, as well as controls for the endogeneity of targeted marketing in inference, all issues that are salient in modern empirical marketing research. We discuss details of the models as well as practical issues involved in translating econometric models of this sort into implementable solutions in the field.

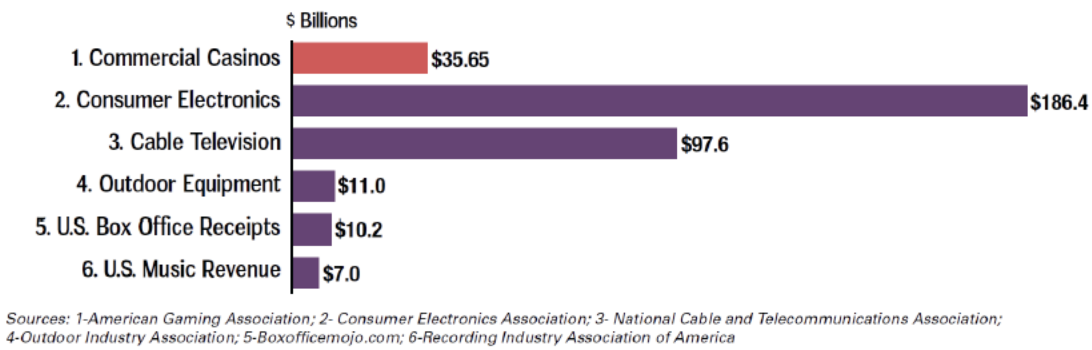
We also propose new ways to accommodate the endogeneity implied by nonrandom historical targeting by the firm in the estimation of empirical models that utilize the firm's transactional data. Our solution is simple to implement, scales easily with large datasets, and exploits the internal information of firms which is usually available in such settings. Our solution involves estimating separate response models for fixed partitions of the historical score variable on which targeting is based on. We expect our approach to be useful in other situations in which researchers work closely with firms and have some knowledge of the firm's past targeting rule. We also discuss generalizable learnings for academics interested in diffusing science to practice that may be relevant in other situations. Our final implementation involves over 120 separate estimated models of consumer behavior (separated by segment, casino and outcomes), over 180+ variables in each model, and over 20,000 parameters estimated across these models. We believe the scale of the model development and implementation, and the evaluation of its impact via a large-scale randomized field experiment is novel to the marketing literature. In our implementation, the ex-post randomized field experiment serves to validate the econometric model we propose, while the quasi-experimental variation induced in the firm's targeting rule serves to identify its key parameters. We see merit in combining models and experiments in this manner, rather than relying on purely model-based or experiment-based approaches to this problem.

Our study adds to a small, burgeoning literature that has used field interventions to assess and demonstrate the validity of econometric marketing models. This includes Mantrala et al. (2006) on pricing aftermarket goods; Cho and Rust (2008) on automobile replacement; Simester et al. (2009) on catalog mailing; and Misra and Nair (2011) on salesforce compensation. Our study also adds to an emerging literature in marketing documenting the applications of Marketing Science models within firms in the real-world (see for instance, Lilien et al. 2013 who review papers associated with the ISMS practice prize). Our work is related to a large literature in marketing that has investigating the value of conditioning promotion allocation on behavioral history (see for instance, Rossi et al., 1996; Ansari and Mela 2003 for representative papers; and Arora et al. 2008 for a review). The most closely related within these are a subset of studies that uses partial or full knowledge of the rules by which marketing is allocated across units to accommodate the reverse causality associated with nonrandom targeting. In this respect, our approach is closest to two papers. First, our method can be thought of as a generalization to a likelihood-based setting of Hartmann, Nair and Narayanan's (2011) Regression Discontinuity based strategy for identifying response under targeting. The advantage of this approach relative to that strategy is our approach utilizes the variation across all consumers within a bin for inference, and is therefore more efficient than the Regression Discontinuity approach, which bases inference on the behavior of only marginal consumers who fall on the edges of targeting bins. Secondly, our method can also be thought of as extending Manchanda, Rossi and Chintagunta's (2004) contribution for handling targeting, to more general behavioral targeting situations where the targeting is a function of the targeted consumer's historical actions. Finally, our work is also related to the literature that has used randomized experiments in field settings to break the endogeneity associated with targeting (e.g., Simester et al. 2009; Goldfarb and Tucker 2011; Sahni et al. 2014).

In 1970, John Little noted with concern that,

> "The big problem with management science models is managers practically never use them. There have been a few applications, of course, but practice is a pallid picture of the promise." (Little 1970).

Figure 1: Comparing U.S. Gaming Revenues to Other Entertainment Spending (AGA, 2012)



See Bucklin and Gupta (1999); Leeflang and Wittink (2000); Roberts (2000); Winer (2000); Lodish (2001); Sinha and Zoltners (2001); Van Bruggen and Wierenga (2001) for various perspectives on the research practice divide. We believe that the availability of large quantities of consumer-level data and the increased recognition of the power of analytics provides for guarded optimism that the trend has turned in the other direction in 2010-s. For example, we are witnessing a rapid diffusion of models built on Marketing Science into practice (e.g., WCAI 2014). We hope our demonstrated improvement from the adoption of an analytics-driven approach to Marketing accelerates this productive collaboration between academic and industry even further.

The rest of the paper discusses the industry context, describes the model framework, discusses the data and results, and presents the results from the field evaluation.

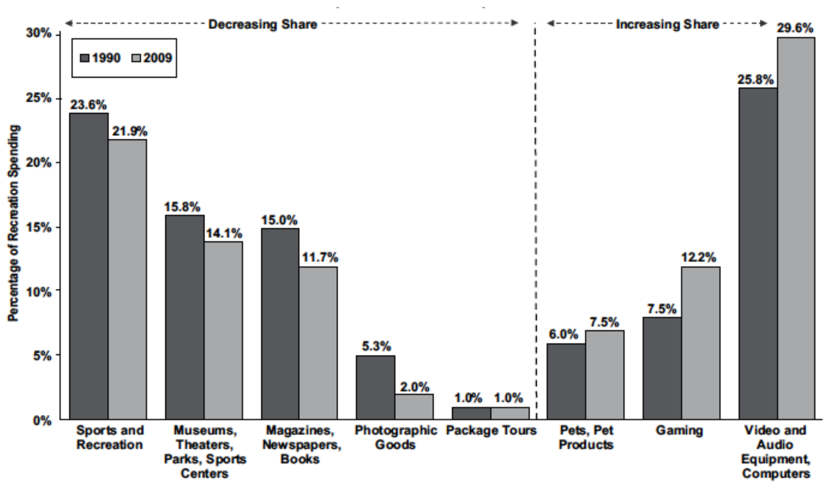
2 Background on Gaming

The market for gaming is part of the hospitality and entertainment sector of the economy. Estimates place the size of the market at about $35.6B in 2011.[1] The market is big. In 2011, the gaming industry employed an estimated 339,098 people who earned $12.9 billion in wages, benefits and tips. Commercial casinos also paid about $7.9 billion to states and localities in the form of direct gaming taxes. Market research by VP Communications, Inc. and pollster Peter D. Hart, reports that more than one-quarter (27%) of the U.S. adult population visited a Casino during 2011, totaling about 59.7 million people (AGA 2012). Figure 1 compares annual consumer spending in commercial casinos to that in a variety of other entertainment channels. Commercial casinos ranked third, ahead of music, outdoor equipment and box office receipts. As American recreational spending rose over the last decade, the share of the spending on gaming grew more quickly than any other component of the recreation sector (see Figure 2; Bazalon et al. 2012). Clearly, gaming is an important part of the entertainment economy.

Within the commercial Casino market, the state of Nevada alone accounts for about 30% of total revenues ($10.7B in 2011). Gambling in Nevada started in the 1800s, but was formally legalized in 1931 as a way of generating jobs in the aftermath of the great depression. Commercial casinos started

[1]Based on commercial Casino revenues reported in AGA (2012). Commercial casinos are profit-making businesses owned by individuals, private companies, or large public corporations. The term "commercial Casino" is used in the United States to indicate a gaming facility that is not owned and operated on Native American lands by a tribal government.

Figure 2: Consumer Spending on Recreation Between 1990 and 2009 (Bazelon et al. 2012)



Source:

[1] "Table 677: Personal Consumption Expenditures by Function," The 2012 Statistical Abstract, U.S. Census Bureau, http://www.census.gov/compendia/statab/cats/income_expenditures_poverty_wealth.html (last accessed October 7, 2011).

Notes:

[1] Based on total recreation expenditure of $314.7 billion in 1990 and $897.1 billion in 2009.

[2] Video and Audio Equipment, Computers includes video and audio equipment; information processing equipment; and services related to video and audio goods and computers.

[3] Sports and Recreations includes sports and recreational vehicles; other sporting and recreational goods; and maintenance and repair of recreational vehicles and sports equipment.

[4] Museums, Theatres, Parks, Sports Centers includes membership clubs and participant sports centers; amusements parks, campgrounds, and related recreational services; admissions to specified spectator amusements; motion picture theaters; live entertainment, excluding sports; spectator sports; and museums and libraries.

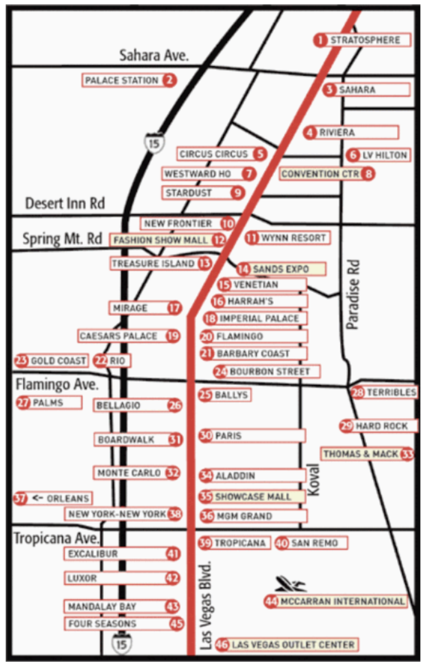
off in Las Vegas, NV in the 1940-s with the opening of El Ranco Vegas, the first Casino, and later, the well-known Flamingo Hotel and Casino started by the mobster, "Bugsy" Siegel. After the U.S. Congress passed the Racketeer Influenced and Corrupt Organizations Act in 1970, the early influence of organized crime in the Casino business reduced and commercial Casino management became more professionalized (Encyclopedia.com, 2012). In 2011, the gaming revenues of the roughly 40+ casinos on the 4-mile stretch of Las Vegas Boulevard known as "the Strip" alone accounted for approximately US$6.1B. This makes this area the top commercial Casino market in the country. Figure 3 depicts the Casinos on the Strip. There is considerable agglomeration. The consolidation helps in demand aggregation, but also results in an intense competitive environment for casinos on the Strip. More recently, casinos on the Strip now face competition from the growth of international markets like Macau, from the gradual relaxation of gambling rules across states within the US, and increasingly from online gambling outlets.

We now discuss some key aspects of casinos that are relevant to understanding the context in which a marketing analytics solution is developed.

Casinos and Product Differentiation

At a broad level, commercial casinos in the U.S. are differentiated in scale and scope into two types, namely destination casino resorts and regional casinos (Bazalon et al. 2012). Destination Casino resorts are large facilities offering gaming, entertainment, retail and convention facilities, and involve development costs that often exceed a billion USD. Destination casino resorts attract visitors from all

Figure 3: Agglomeration of Commercial Casinos on the Las Vegas "Strip"



over the world. Regional casinos are smaller operations, catering mostly to customers within driving distance, and focused primarily on gaming. The mix of destination versus regional casinos in a location is determined by a variety of regulatory, demand and competitive factors. Most of the casinos on the "Strip" tend to be destination casinos.

Destination casinos are multi-product firms providing bundles of entertainment, lodging, retail and gambling options to consumers. A key feature is complementarities in demand across offerings. Good lodging, entertainment, and food and beverage (henceforth F&B) options attracts patrons, who in turn stay longer and spend more on recreational activities. Consequently, casinos often implement loss-leader pricing on several offerings, particularly on lodging and F&B, in combination with targeted price and promotion discrimination. Casinos target high value consumers with subsides on stays and perks and offset these promotional costs with the option value of increased spending by the targeted consumers. Identifying such consumers and finding the right mix of promotions to offer them then becomes a key component of the business model of the firm. In 2010, commercial casinos in the US earned about 69% of their revenues from gaming, 13.2% from F&B, 10.4% from Hotel and lodging and the remaining 7.1% from other activities (e.g. golf, spa, concerts).

Marketing and the Challenge of Targeting

The proliferation of gaming outlets as well as the agglomeration of several competing options on locations like the "Strip" implies competition for consumers is intense. Hence, marketing becomes very important for driving revenue. Casinos offer a variety of promotions to consumers including discounted rooms and entertainment tickets, credits for subsidized play at slots and tables (referred to as "free-play"), discounts on food/drinks, as well as concierge service, subsidized credit and risk-sharing agreements to high-spending "high-roller" consumers.[2] These offers or "comps" − short for "complementary" − are marketed via a variety of channels including direct-mail, email, online advertising, and banners. Much of marketing effort is targeted. As a general rule, more attractive promotions are offered to those that are expected to play more.

Targeting in the gaming context is a complicated problem. The extent of consumer heterogeneity is huge, which complicates the task of determining the consumers with the highest marginal propensity to respond to a promotion. Casinos face the task of simultaneously attracting high-spending consumers while avoiding highly skilled "experts" who win back from the house more than they wager. Casinos would also like to avoid consumers who utilize comps but do not play at the resort. They would also like to avoid consumers who wager nothing more than their free-play dollars, thereby gaining the upside from the promotion, with little downside for themselves and no gain for the "house." Unfortunately, it is not easy to sort out desirable consumers from undesirable ones based on observed socio-demographic characteristics, leading to a difficult adverse selection problem. Casinos attempt to solve some of these difficulties by using history-dependent marketing policies, targeting offers based on functions of a consumer's observed past play behavior (more on this below). Unfortunately, application of this policy over time has caused consumer expectations to adjust. Many consumers now expect free-play and comps to be automatically allocated when they spend more, and may even stop patronizing a casino if it does not offer them significant comping. Consequently, comp-activity and promotional spending in Vegas casinos has grown significantly in recent years, and many industry observers feel that much of comping does not drive incremental demand, being delivered to many without measurable incremental effect on spending. While in the past, comping was seen as a reward that had the causal effect of increasing play spending, now, many believe some consumers see it as a pre-condition to spending. Past comping has created in effect, a "comp-sensitive" segment, a form of moral-hazard caused by targeted marketing policies. In addition, when casinos that promote more also attract more "comp-sensitive" consumers, the adverse selection problem is also deepened. Both issues accentuate the difficulty of targeting and optimizing marketing effort in this setting. Moreover, there is also an overarching concern that targeting more promotions to those who have played a lot in the past may be ineffective, because those consumers may already be on the flat or declining part of their promotion response curve. The history-based allocation may then be targeting promotions to those who would have played anyway, which ends up losing money. These issues have parallels to issues faced by firms in other industries in managing their long-run promotion policies (e.g., manufacturers offering automobile promotions to car-buyers, retail sales to apparel consumers, and trade-promotions to retailers have been concerned that promotions end up losing money due to analogous reasons to above).

A second complication is finding the right match between promotions and consumer preferences. Different consumers have different preferences over hotel rooms, F&B or free-play offers. An ideal policy will target a mix of promotions to each consumer based on what produces maximal marginal benefit at minimal cost. This requires a disaggregate model of heterogeneous consumer tastes that can be made the basis of individual-level policy making. Many casinos lack such sophisticated analytics capabilities. While casinos have made significant progress in identifying cheats using individual-level data, much of their analytics are based on RFM (Recency-Frequency-Monetary value) models, that preclude more finer segmentation. Casinos also offer promotions in packages, bundling varying levels

[2] Examples of risk-sharing for high-rollers include returning a negotiated percentage of losses back to the consumer.

of differing offers into tailored packages. These packages are often offered concurrently with component promotions. Finding the right match between a consumer and a bundle or component of promotional options is thus a large-scale combinatorial, mixed bundling problem.

A third complication is that many destination casinos own more than one property, each of which may run marketing campaigns in parallel. For instance, the MGM group manages 11 casinos in Las Vegas. In this situation, it is possible that promotions cannibalize demand within the product portfolio.[3] Targeting in this situation has to be co-ordinated such that a consumer is not unprofitably attracted away from a high-margin, high-spend property to a low-margin, low-spend one. Preventing trading-down of this manner requires understanding consumer preferences not just over promotions, but over promotion-property combinations, so the targeted promotion incentivizes the focal customer to self-select into the preferred property from the firm's perspective. Further, there is a need to understand the impact of promotions at the property level due to the need for good demand forecasting. Lodging and F&B are capacity-constrained resources, and accurate forecasting of the expected visitation and utilization of these resources in response to campaigns is important for effective operational managing and planning.

Finally, promotion management is not a static problem. Consumers exhibit significant state-dependence and persistence in their visitation and play behavior. Thus, current promotions have long-lived effects into the future, by affecting the stickiness and profile of repeat business. Incorporating these dynamic effects of promotions is important to get an accurate picture of the ROI profile from the promotions, and to allocate them appropriately based on their expected long-run benefit to the firm.

Current Practice

Some aspects of current practice in targeting has been alluded to above. Many casinos are not analytically sophisticated in their marketing targeting practice, and employ even more crude, heuristic-based targeting rules compared to empirically-driven history-based strategy. Targeting practice at MGM prior to implementation of the new analytics solution described here was more sophisticated than at many other casinos, but subject to several of the concerns outlined above. Like many casinos, MGM's practice involved use of a specific form of history-based targeting. To understand the rule, it is useful to define a few metrics commonly used in the casino setting.

- *Coin-in:* is the total dollar outlay by a consumer at a play occasion.
- *Hold Percentage*: is interpreted as the long-run average return for the casino when the consumer plays a dollar in repeated plays. For example, if a consumer bets \$1 at a slot machine, and the casino has programmed the machine such that it returns \$0.8 to the consumer on average, the *Coin-in* is \$1, and the *Hold Percentage* is = 20%.
- *Theoretical Win* or "*Theo*": is widely used in casino mathematics as a measure of how much money a casino is *expected* to win from a consumer on a given play. It is defined as Coin-in $\times$ Hold Percentage for the play. It is different from the *Actual Win*, another common metric, because actual outcomes may be influenced by random factors like the realization of the play's

[3]For example, a smart consumer may utilize a corporate lodging promotion to stay at a property in a casino chain, and spend his entire visit availing of concurrent property-specific promotions at other properties within the chain, with no incremental spending realized from the visit to the casino.

hold. Essentially, how much money a casino can make from a consumer play is a random variable. The Actual Win is the realization of that random variable, and the Theo is the expected value of that random variable. For example, if a player plays \$100 on Slot Machine A, which has a hold percentage of 20%, and then later that day plays \$100 on Slot Machine B, which has a hold percentage of 15%, the *theoretical win* from the casino's perspective for that player for that day is $\$100 \times .20 + \$100 \times .15 = \$35$. Actual win may be different from \$35 because slot machine A kept \$21 (and not the expected value of \$20) for the casino on the consumer's play there, and slot machine B kept say \$13.5 (and not the expected value of \$15) when the consumer played there. Thus, Actual Win over the day $= \$21 + \$13.5 = \$34.5$. The Average Daily Theoretical (or ADT) is simply the average theo over all the individual games played by the consumer over the last N months of a consumer's visits, where N varies depending on the casinos ability to store and manage consumer data. The Average Daily Actual is defined analogously.

MGM, like other casinos, tracks traditional gaming industry metrics like actual win and theoretical win for its customers. Promotions are allocated based on bins of average theo and actual wins on the observed trips over the previous N months by the consumer (we cannot reveal the exact value of N due to business confidentiality concerns). In practice, theo and actual wins are highly correlated across consumers, hence, one can think of this as segmentation on RFM criteria linked to average theo. More promotions are allocated to those that are observed to have higher realized theo win in the trips over the last year. Once consumers are scored on the average theo + demographics, a marketing campaign involving a specific set of promotions is considered. Those with the highest scores get the most attractive promotions, those with smaller scores get the less attractive ones, and so forth, where "attractiveness" of a promotion is assessed based on managerial judgment and knowledge. The bulk of the promotions are targeted directly to consumers via direct-mail and/or email.

Opportunity for Analytics to make an Impact

Analytics can improve the above targeting rule significantly. Casinos are a data-rich environment. Due to large investments in data-warehousing technologies, and due to the wide-spread adoption and use of loyalty cards, most transactions made by a consumer during a visit to any of the portfolio properties of MGM are tracked.[4] These data can be used to build detailed models of consumer behavior and consumer response to promotions. These facilitate development of model-based metrics of consumer value, which can be utilized for subsequent targeting.

Scoring consumers' value on the basis of their average theo over recently observed trips has several disadvantages. First, it induces large variability in a given consumer's value across trips that is driven by random factors outside the consumer's control or unrelated to his preferences. The variability implies valuable consumers may drop in and out of campaigns and are not consistently targeted. Second, it does not help understand how promotions drive value, for instance, by understanding whether promotions work by increasing visitation, or by changing the property chosen conditional on visitation, or by changing spending conditional on property choice and visitation. Understanding these may be important for formulating and fine-tuning marketing strategy. Third, it does not provide a forward-looking measure of value that assesses the extent to which a consumer is likely to be profitable

[4]Cash is first exchanged for a play-card linked to a unique loyalty-card ID or for chips on the casino floor. Most aspects of subsequently play (where, when, how long and how much played), as well as activities (rooms stayed at, shows watched), and promotions allocated are thus captured in the database.

to MGM in the *long-run*. For instance, a consumer may have wagered little in his first visit due to a trip-specific reason, but may yet be profitable in the long-run to the casino because his base propensity to spend at the firm is high. Conditioning value on a consumer's recent trip outcomes misses this component of value. Model-based metrics can address these disadvantages.

Our model-based metric has the advantage that it uses data on observed behavior at all past visits (and not just the most recent visits) to measure customer value. Hence, it is less variable than recent-trip metrics. Additionally, for consumers on which very little data exist, the model pools information from the behavior of similar consumers to provide a less noisy estimate of value compared to using only recent trip information. It also uses information across the entire range of activities by the consumer to measure how promotions affect behavior. Moreover, model-based metrics are both history-dependent (retrospective) and forward-looking (prospective). In the example above of the customer who visited once but spent little, the model based metric will use the first-visit information on the consumer *in conjunction* with the observed long-run spending of *other similar* consumers. Suppose it turns out in the data that these other consumers spend highly in future visits even though they spent little on their first. The model will then identify the focal consumer in the example as profitable in the long-run and a viable candidate for targeting, even though his observed first-trip spending was low. Finally, by modeling consumer behavior across the full product-line, models that pool data across properties enables better assessment and management of cannibalization within the firm's product portfolio.

A second area where analytics can make an impact is to improve the match between the consumer and the promotion bundle. Models estimated on the data predict the expected marginal response of each consumer type for each combination of offers that make up hypothetical promotion bundles. Thus, they provide a *promotion-bundle*−specific score for each customer. In parallel, advances in computing power enable one to search for the optimum bundle for each consumer taking these model-predicted responses as input. Together, this enables customizing promotions to each customer and facilitates development of a scalable, data-driven micro-targeting policy. This is in essence the new approach implemented at MGM.

In the reminder of the note, we describe the features of this approach, details on the underlying econometrics, and report on results from its field evaluation.

3 Model Framework

The goal of the empirical model is to deliver predictions of the following for consumer i in month t.

1. Whether i will visit one of the MGM casinos in month t. Denote this by the indicator variable $y_{i0t}^{(1)}$.

2. Conditional on visiting, whether i visits property $j \in (1, .., J)$. Denote this by the indicator variable $y_{ijt}^{(1)}$.

3. Conditional on visiting, how much will i spend. Denote this by the continuous variable $y_{it}^{(2)}$.

Collect these in a vector $\mathbf{y}_{it} = \left(y_{i0t}^{(1)}, y_{i1t}^{(1)}, .., y_{iJt}^{(1)}, y_{it}^{(2)}\right)$. Assume that there are K_{jt} different bundles of marketing promotions offered at property j in month t and let x_{ikjt} be an indicator for whether the k^{th} promotion bundle was offered to consumer i for utilization in property j in month t. A promotion bundle is a particular combination of offers valid at one or more properties at the casino (e.g., 2 Tickets to a show + \$100 free-play valid only at the Bellagio; or suite upgrade at any of the MGM properties). Collect the promotional offers valid for property j for month t in a vector $\vec{x}_{ijt} = \left(x_{i1jt}, .., x_{ikjt}, .., x_{iK_{jt}jt}\right)$, and collect all the promotion vector across properties for the individual in an array $\mathbf{x}_{it} = (\vec{x}_{i1t}, .., \vec{x}_{ijt}, .., \vec{x}_{iJt})$. Let d_i be a vector containing the observed socio-demographics of consumer i. Our general model is of the form,

$$\mathbf{y}_{it} = f\left(\mathbf{x}_{i,t-\tau:t}, \mathbf{y}_{i,t-\tau:t-1}, d_i, \boldsymbol{\epsilon}_{it}; , \Omega_i\right) \tag{1}$$

where, $f(.)$ is a parametrically chosen link function (discussed below), and $\boldsymbol{\epsilon}_{it}$ is a vector of consumer and month specific unobservables that are observed by the consumer and incorporated into his decision making, but unobserved by the econometrician. Equation (1) allows for state dependence in consumer behavior by allowing current actions to depend on past outcomes over the past τ periods. Equation (1) also allows for heterogeneity in consumer response because the model parameters, Ω_i, are allowed to be consumer specific. The goal of inference is to use the data to estimate the parameters Ω_i. The subset of Ω_i relating to the direct effect of $\mathbf{x}_{it}$ on $\mathbf{y}_{it}$ represent the causal effect of promotions on outcomes, and are key to identifying a set of desirable consumers for subsequent targeting. The data for estimation includes observations on $(\boldsymbol{y}_{it}, \boldsymbol{x}_{it}, \boldsymbol{d}_i)$ for a large sample of consumers (over 1M) over a roughly two year horizon, during which every visit of each i to MGM is tracked along with every promotion offered.

We now discuss the specifications we choose for $f(.)$ for estimation.

3.1 Nested Logit Model of Visit and Property Choice

We model the discrete-choice of whether to visit the casino in a given month, $y^{(1)}_{i0t}$, and the choice of which property to visit, $y^{(1)}_{ijt}$, as a nested logit model. To operationalize the model, we need to accommodate the fact that the consumer also faces a discrete choice over use of a promotion bundle conditional on visit to a property.[5] The self-selection of consumers into a promotion bundle is in and of itself informative of types (e.g., Chiang 1995), and we would like to accommodate the information content of these choices into our estimation procedure. To accommodate this aspect, we specify the lower-most nest of the discrete-choice model as a choice over use of one (or none) of the offered promotion bundles. The higher level nests then captures the choice of property, or to not visit. Figure (4) depicts the nesting structure.

We specify the probability that consumer i chooses bundle k at property j in month t, ϱ_{ikjt}, as,

$$\varrho_{ikjt} = \frac{\exp\left(\psi_{ik}\right)}{1 + \sum_{k=1}^{K_{jt}} \exp\left(\psi_{ik}\right)} \tag{2}$$

where, $\psi_{i.}$ are bundle-specific parameters. The probability of visiting property j without using any of the offered bundles is, $1 - \sum_{k=1}^{K_{jt}} \varrho_{ikjt}$.

At the second level of the nest, we specify the probability of visiting property j as $\Pr(y^{(1)}_{ijt} = 1) = \frac{\exp(v_{ijt})}{1+\sum_{j=1}^{J}\exp(v_{ijt})}$, where the consumer-specific attractiveness of property j in month t, v_{ijt}, is specified as,

$$v_{ijt} = \varsigma^{(1)}_{ij} + g\left(\mathbf{y}_{i,t-\tau_1:t-1}, \mathbf{x}_{i,t-\tau_1:t}, d_i; \varsigma^{(2)}_{ij}\right) + \sigma_j \ln\left[1 + \sum_{k=1}^{K_{jt}} \exp\left(\psi_{ik}\right)\right] \tag{3}$$

[5]The casino allows consumers to use only one offer bundle per visit.

In the specification above, $\sigma_j \in (0,1)$ is a property-specific parameter that captures the effect of the promotions offered on a consumer's utility of visiting a property. σ_j serves as a weight on the "log-sum" for the lower nest representing the expected utility from utilization of the most preferred promotion bundle for property a j. $g(.)$ is a function of the past τ_1 trips made by the consumer which we use to allow for state dependence in demand in choices. We specify $g(.)$ to be linear in main and interaction effects of past visitation behavior, promotion utilization and demographics, and indexed by property-specific parameter vector $\varsigma_{ij}^{(2)}$. Allowing for $g(.)$ helps improve fit and capture heterogeneity. Finally, $\varsigma_{ij}^{(1)}$ is a property-j specific intercept. The probability of not visiting any of the MGM properties in a month t is by construction, $\Pr(y_{i0t}^{(1)}=1) = 1 - \sum_{j=1}^{J} \Pr(y_{ijt}^{(1)}=1)$.

3.2 Log-linear Model of Spending

We model spending conditional on visit and property choice as a "Burr" model,

$$y_{it}^{(2)} = \mu \left[\frac{\exp\left(h\left(\mathbf{y}_{i,t-\tau_2:t-1}, \mathbf{x}_{i,t-\tau_2:t}, d_i; \theta_i\right)\right)}{1+\sum_{k=1}^{K_{jt}} \exp\left(h\left(\mathbf{y}_{i,t-\tau_2:t-1}, \mathbf{x}_{i,t-\tau_2:t}, d_i; \theta_i\right)\right)} \right]^{1/2}$$

In the above specification, $h(.)$ is a function of the current and past τ_2 trips made by the consumer which allows for state dependence in spending. We allow $h(.)$ to be a flexible linear function comprising of main and interaction effects of current and past visitation behavior, promotion utilization and demographics, indexed by the parameter vector θ_i. μ is a saturation parameter that puts an upper bound on predicted spending. We set μ to be 1.5 $\times$ the maximum observed per-trip spending across consumers. The Burr model above allows expenditure to be positive and bounded and prevents the model from predicting unreasonably large values of spending in prediction settings. Thus, under this model, one interprets the observed spending as a flexible fraction of the maximum spend, \$$\mu$.

We now collect the set of parameters to be estimated in $\Omega_i \equiv (\{\psi_{ij}, \varsigma_{ij}^{(1)}, \varsigma_{ij}^{(2)}, \sigma_j\}_{j=1}^{J}, \theta_i)$.

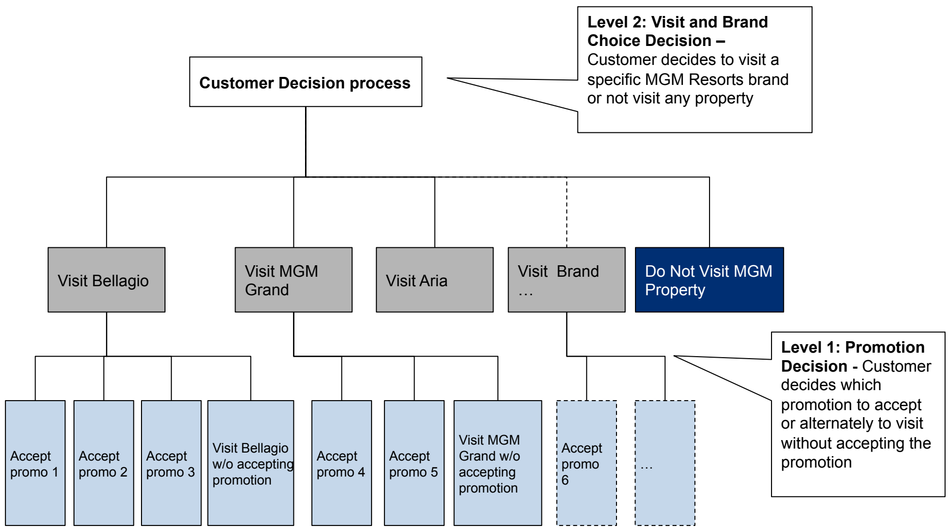
4 Estimation

We estimate all the models presented above by maximum likelihood. Before discussing specific details, we first discuss how we address the endogeneity concern that arises due to the history-dependent nature of the targeting rule used by MGM under which the data were generated.

4.1 Endogeneity Concern

The concern about endogeneity becomes relevant because we are interested in using the model for segmenting consumers and not just for prediction. While estimated parameters need not have a causal interpretation for purely predictive purposes, segmentation requires understanding the causal effect of promotions for each customer. Unfortunately, because more promotions were targeted in the data to consumers who play more, *a priori* we cannot say whether any positive covariation we find in the data between outcomes and promotions is the result of a causal effect of those promotions on outcomes, or the effect of outcomes on promotion allocation as induced by the targeting rule. This is the identification problem inherent in the analysis. In our setting, we find accommodating the effect of the reverse causality induced by targeting critical to avoid overstating the effect of promotions.

Figure 4: Nesting Structure Used in Model Setup



Our approach uses a partial (but not perfect) solution to the problem. To understand the philosophy behind our approach, note that our goal in addressing the endogeneity is not necessarily to measure without bias a specific coefficient with an economic interpretation for a specific marketing variable (e.g., the elasticity). Our concern is less about consistency of coefficients and more about producing a better out-of-sample prediction of promotional effectiveness which could ultimately be a reliable input to the formulation of an improved promotion policy. We will demonstrate later in the paper via a field implementation that our approach produces a better promotion strategy for the firm than the existing policy (or doing nothing).

To understand our approach, we return to the notation used in Equation (1), where the model is set up in general terms relating outcomes, $\mathbf{y}_{it}$ to $(\mathbf{x}_{i,t-\tau:t}, \mathbf{y}_{i,t-\tau:t-1}, d_i, \boldsymbol{\epsilon}_{it})$. Assume for a moment that promotions $\mathbf{x}_{it}$ are randomly allocated to each agent i. Then, ignoring the initial condition, we can write the likelihood across agents as,

$$\mathcal{L}\left(\{\Omega_i\}, \phi\right) = \prod_{i=1}^{N} \prod_{t=\tau+1}^{T_i} f_{y_t|y_{t-\tau:t-1}, \boldsymbol{x}_{t-\tau:t}, d}\left(y_{it}|y_{i,t-\tau:t-1}, \mathbf{x}_{i,t-\tau:t}, d_i; \Omega_i\right) \prod_{i=1}^{N} \prod_{t=\tau+1}^{T_i} f_x\left(\mathbf{x}_{it}; \phi\right) \quad (4)$$

where, $f_{y_t|y_{t-\tau:t-1}, \boldsymbol{x}_{t-\tau:t}, d}(.)$ is the conditional density induced on $\boldsymbol{y}$ by $\boldsymbol{\epsilon}$. $f_x\left(x_{it}; \phi\right)$ is the density of $\mathbf{x}_{it}$, and ϕ are parameters of that density. The key to note in this situation is that the likelihood factors in $\mathbf{x}$: the density of $\mathbf{x}$ across consumers is not informative about the underlying parameters of interest, Ω_i, because the variation of $\mathbf{x}$ across consumers is not conditioned on Ω_i. Since the density contribution of $\mathbf{x}$ is not a function of Ω_i, it can be ignored when searching for Ω_i that maximizes the likelihood of the data.

On the other hand, when promotions are targeted to consumers based on their behavior, we should write the conditional likelihood as,

$$\mathcal{L}\left(\{\Omega_i\},\phi\right)=\prod_{i=1}^{N}\prod_{t=\tau+1}^{T_i}\left[f_{y_t|y_{t-\tau:t-1},\boldsymbol{x}_{t-\tau:t},d}\left(y_{it}|y_{i,t-\tau:t-1},\mathbf{x}_{i,t-\tau:t},d_i;\Omega_i\right)f_{x|d,\mathcal{H}}\left(\mathbf{x}_{it}|d_i,\mathcal{H}_{it};\Omega_i,\phi\right)\right] \tag{5}$$

The likelihood no longer factors because $\mathbf{x}$ is set with some knowledge of the consumer's type Ω_i, his history $\mathcal{H}_{it}$ and characteristics d_i. Intuitively, now, the variation of $\mathbf{x}$ across individuals is also informative about Ω_i. For instance, the fact that an individual is observed to have a high level of marketing targeted to him in the data now tells the model that he is a "high"-Ω_i type. In this situation, we can no longer ignore the likelihood contribution associated with the density of $\mathbf{x}$. Ignoring that will misspecify the likelihood for Ω_i causing a first-order bias. Moreover, not knowing the true density of $f_{x|d,\mathcal{H}}(.)$ also has the potential to cause a bias, arising from a second-order source of misspecification associated with imposing the wrong density. Hence, to recover Ω_i, the likelihood needs to be augmented with the correct conditional density of $\mathbf{x}$. This logic is similar to that in Manchanda, Rossi and Chintagunta's (2004) analysis, but with the difference that their specification does not allow the density of $\mathbf{x}$ to condition on the consumer's history.

Our Approach Our approach is facilitated by that fact that we know the exact variables on which targeting by the casino is based on namely, average *Theo* and Demographics, and the fact that we observe these variables in the data. Let z_{it} denote the average *Theo* of the consumer over his observed trips to the casino over the previous N months, evaluated at the beginning of period t. Let the subset of demographics used by MGM for targeting be denoted $\tilde{d}_i$. Both z_{it} and $\tilde{d}_i$ are observed in the data. We know that $\mathbf{x}$ depends on Ω_i only through $(z_{it},\tilde{d}_i)$; thus, we can write,

$$f_{x|d,\mathcal{H}}\left(\mathbf{x}_{it}|d_i,\mathcal{H}_{it};\Omega_i,\phi\right)=\underbrace{f_{x|z,\tilde{d}}\left(\mathbf{x}_{it}|z_{it},\tilde{d}_i;\phi\right)}_{\text{part } \mathbb{I}}\times\underbrace{f_{z|\mathcal{H}}\left(z_{it}|\mathcal{H}_{it};\psi\right)}_{\text{part } \mathbb{II}} \tag{6}$$

The likelihood has two parts, the first representing the conditional distribution of $x|z,\tilde{d}$, and the second, the distribution of z given the agent's behavioral history. The first represents the process by which behavioral targeting is implemented given the "score" variable z, and the second represents the process that generates the score. We discuss these in sequence, explaining the challenges we face in characterizing these exactly. We then discuss the econometric procedure we use that circumvents these difficulties.

Part I: Conditional density of $x|z,\tilde{d}$

Part I of the likelihood tells us that we should exploit only the variation in $\mathbf{x}$ *holding* $z_{it},\tilde{d}_i$ *fixed* to learn about the direct effect of $\mathbf{x}$ on $\mathbf{y}$. This part of the likelihood is key to the control for targeting. Intuitively, as $\mathbf{x}$ changes, it produces both a direct effect on $\mathbf{y}$ due to the impact of promotions on outcomes, as well as an indirect effect by changing the set of individuals targeted. Only the first type of variation is useful to measuring the causal effect of $\mathbf{x}$ on $\mathbf{y}$; the second measures the selection induced by targeting. Including the conditional density of $\mathbf{x}$ into the likelihood tells the model that all selection of types that arises from changes in $\mathbf{x}$ happen *only through changes* in z and $\tilde{d}$. Hence, any changes in $\mathbf{y}$ that is associated with changes in $\mathbf{x}$ holding z and $\tilde{d}$ fixed, is useful to learn about the direct effect of $\mathbf{x}$ on $\mathbf{y}$. Conditioning in this manner helps the model utilize that variation correctly.

If we knew the density $f_{x|z,\tilde{d}}(.)$ on the right hand side of (6) perfectly, we could plug it into Equation (5) to handle this aspect. A concern arises from the fact that we do not know $f_{x|z,\tilde{d}}(.)$ perfectly, because the exact targeting function mapping $(z_{it}, \tilde{d}_i)$ to $\mathbf{x}$ is not well documented within the company. The company's promotions are determined for each campaign by an internal committee. The committee decides which $\mathbf{x}$-bundle to target to consumers of a given $(z_{it}, \tilde{d}_i)$ based on business priorities prevalent at the time of design of the promotional campaign. For instance, when determining promotion allocation across customers, MGM may decide it wants to increase visitation at a particular casino, and provide more promotions for that property to high-*Theo* consumers; alternatively, the committee may decide its campaign goal is to increase visitation at the slots, and allocate more slot-specific promotions to high-*Theo* customers. While we know that promotions $\mathbf{x}$ are allocated on the basis of $(z_{it}, \tilde{d}_i)$, we do not have a way of modeling which $\mathbf{x}$ and how much of it will be allocated for any given value of $(z_{it}, \tilde{d}_i)$. Thus, we cannot credibly characterize the conditional density $f_{x|z,\tilde{d}}(.)$. This implies we need a way to conduct inference without knowing the assignment probability.

Analogous situations are often faced by researchers. In many contexts, researchers may know that promotions are assigned on the basis of well-defined segments, but do not know exactly how the firm made the decision to offer a particular offer/creative to a given segment.

Part II: Density of $z|\mathcal{H}$

To understand part II we need to evaluate the process generating z. Recall that z_{it} represents the average *Theo* of the consumer over his observed trips to the casino over the previous N months. However, note that *Theo* is a metric that is specific to an entertainment option, and not to a trip as a whole. *Theo* is calculated as the money spent by the customer on an entertainment option, multiplied by the hold percentage for that option as fixed by the casino. Thus, the z_{it} variable observed in the data is a function of the vector of expenditure outlays by the consumer over the past N trips, denoted $\boldsymbol{y}^{(2)}_{i,t-N:t}$; the vector of expenditure *splits* of that outlay across the various available entertainment options on those past trips, denoted $\boldsymbol{w}_{i,t-N:t}$; as well as the vector of hold-percentages for the various entertainment options available over those trips, $\Gamma_{t-N:t}$ (note, the hold-percentage is not consumer-specific). Thus, we can write,

$$z_{it} = z\left(\boldsymbol{y}^{(2)}_{i,t-N:t}, \boldsymbol{w}_{i,t-N:t}, \boldsymbol{\Gamma}_{t-N:t}; \Omega_i\right) \tag{7}$$

The key issue is that lagged expenditures and expenditure splits ($\boldsymbol{y}^{(2)}_{i,t-N:t}$ and $\boldsymbol{w}_{i,t-N:t}$) are both a function of Ω_i, because they are chosen by the consumer. Hence, z_{it} is itself a function of Ω_i as noted explicitly in Equation (7). One solution to completing the likelihood is to model how the score is generated, by explicitly modeling the RHS of Equation (7). Unfortunately, this is difficult. While we do model total expenditures $\boldsymbol{y}^{(2)}_{i,t-N:t}$, modeling the expenditure *split* across various entertainment options, $\boldsymbol{w}_{i,t-N:t}$, will require us to write down models of how and why a consumer chooses a particular sequence of casino entertainment options and associated expenditure decisions. This is complicated because the expenditure decisions are interrelated (due to a common budget constraint), are potentially driven by state dependence (e.g., Narayanan and Manchanda 2012), and require high-frequency within-trip play data to model. Modeling $\boldsymbol{w}_{i,t-N:t}$ in this manner is beyond the scope of this project. Hence, we need a way to conduct inference without having to know the density of z.

Analogous situations to this problem are often faced by researchers. In many situations, researchers may know that promotions are assigned on the basis of a scoring variable that is observed in the data. When the score is informative of the response parameters, its likelihood contribution cannot be ignored. However, researchers may be unable to model the data generating process for the scoring variable credibly, because often they do not know all determinants of the score. Treating the unknown determinants of the score as random noise is not a solution either. In behavioral targeting situations, it is highly likely that some of these unknown (or un-modeled, like in our situation) determinants of the score reflect historical actions taken by the customer. Hence, these determinants of the score are correlated with the response parameters. For instance, in the above situation, the unknown determinants of the score, $\boldsymbol{w}_{i,t-N:t}$, reflect historical play behavior and are a function of Ω_i (and are thus correlated with Ω_i). Because of this reason, we cannot ignore the contribution of $f_{z|\mathcal{H}}(.)$ to the full likelihood for estimating Ω_i.[6]

Recap of Econometric Problem To recap the discussion so far, we can combine Equations (5), (6), and 7), to note that the full likelihood that includes behavioral targeting is,

$$\begin{aligned} \mathcal{L}\left(\{\Omega_i\},\phi,\psi\right) = \quad & \textstyle\prod_{i=1}^{N_r}\prod_{t=\tau+1}^{T_i} f_{y_t|y_{t-\tau:t-1},\boldsymbol{x}_{t-\tau:t},d}\left(y_{it}|y_{i,t-\tau:t-1},\mathbf{x}_{i,t-\tau:t},d_i;\Omega_i\right) \\ \times \quad & f_{x|z,\tilde{d}}\left(\mathbf{x}_{it}|z(\boldsymbol{y}^{(2)}_{i,t-N:t},\boldsymbol{w}_{i,t-N:t},\mathbf{\Gamma}_{t-N:t};\Omega_i),\tilde{d}_i;\phi\right) \\ \times \quad & f_{z|\mathcal{H}}\left(z(\boldsymbol{y}^{(2)}_{i,t-N:t},\boldsymbol{w}_{i,t-N:t},\mathbf{\Gamma}_{t-N:t};\Omega_i);\psi\right) \end{aligned} \tag{8}$$

We can summarize the econometric difficulty as two-fold. (a) the last two terms in the likelihood are functions of Ω_i and cannot be ignored. At the same time, (b), the researcher is unable to model these explicitly, either because of lack of knowledge of the process or due to incomplete information. Further, the missing information is not random.

Our Solution Our solution to the issue is based on exploiting additional knowledge of the targeting rule. In particular, targeting at the firm is implemented via a discrete form of segmentation. The firm bins the *Theo* z_{it} into many buckets and combines these buckets with $\tilde{d}_i$ to construct segments. Promotion assignment is based on these segments. For each campaign, the firm's committee decides which segments should be considered for each possible promotion bundle. Because consumers are heterogenous, and promotions are multi-dimensional, its not necessarily the case that consumers in the highest *Theo* segments are allocated the most attractive promotions i.e., bins are not ordered on the basis of the attractiveness of possible promotion bundles. Once the committee decides which segment is eligible for a chosen promotion bundle, a *subset* of consumers within that segment are chosen to be sent the promotion. The reason for sending the promotion only to a subset within the segment rather than to all within it, is the firm faces margin or cost constraints that prevent it from blanketing everyone in the segment. For instance, the committee may face a dollar cap on the total promotional spending at its disposal in a given campaign; or it may face a limit on the number of

[6] Zantedeschi et al. (2014) point out that the density of how promotions are targeted to individuals can be ignored for inference of Ω_i if targeting is purely a function of modeled history (i.e.., if the score z is a function of past $\boldsymbol{y}$ and $\boldsymbol{x}$ alone). This simplification does not obtain when there are unobservables (like $\boldsymbol{w}_{i,t-N:t}$) that drive the score and are correlated with the response parameters; or if the targeting is based on response parameters directly over and above the history.

promotions it can offer for stays at a particular property or for visits at a particular show due to capacity constraints at those locations.

In most campaigns, a randomly picked subset of consumers within the segment are then assigned the promotion bundle. In a smaller set of campaigns, the random sampling is done after picking a subset of consumers within the segment whose spending in the previous trip crossed a given amount.[7] A consequence of this assignment scheme is that, conditional on past trip spending, whether a consumer in a bin *receives* the promotional bundle is essentially random. It is also the case that different properties face different margin or cost constraints based on their priorities and competitive situation, and this generates variation in the number of consumers picked within each bin. Both sources of within-segment variation in promotions are not correlated with consumer tastes. We explain below how we can exploit this design for econometric inference.

Our method works as follows. First, we divide z into discrete bins and form segments by combining the bins of z and bins of demographics used by MGM for targeting. Formally, letting $i_z \in (1, .., \mathcal{I}_z)$ denote the bins on the z dimension, and $i_{\tilde{d}} \in \left(1, .., \mathcal{I}_{\tilde{d}}\right)$ denote the bins on the demographics dimension, we define $R = \mathcal{I}_z \times \mathcal{I}_{\tilde{d}}$ segments corresponding to each combination of i_z and $i_{\tilde{d}}$. Because the R segments are defined on the basis of observables, we can a priori assign all observations to one of the R segments. We then estimate a separate model for each such segment $r \in (1, .., R)$, to estimate a segment-specific parameter vector, $\Omega^{(r)}$. This controls for the endogeneity of targeting.

To understand why the approach works, suppose we focus only on bin r, in which z takes values between $(\underline{z}_r, \overline{z}_r)$. Consider the subset of N_r consumers who have been assigned *a priori* to segment r, and let $\Omega^{(r)}$ represents the segment-r specific parameter vector. Consider the likelihood for only this segment,

$$\begin{aligned} \mathcal{L}\left(\Omega^{(r)}, \phi\right) \quad = \quad & \prod_{i=1}^{N_r} \prod_{t=\tau+1}^{T_i} f_{y_t|y_{t-\tau:t-1}, \boldsymbol{x}_{t-\tau:t}, d}\left(y_{it}|y_{i,t-\tau:t-1}, \mathbf{x}_{i,t-\tau:t}, d_i; \Omega^{(r)}\right) \\ & \times \qquad f_{x|\tilde{d}}^{(r)}\left(\mathbf{x}_{it}|\tilde{d}_i; \phi\right) \end{aligned} \tag{9}$$

The key point to note is the variation of $\mathbf{x}$ *within* $z \in (\underline{z}_r, \overline{z}_r)$ does not depend on $\Omega^{(r)}$, because who gets assigned the promotion within the segment is random. Hence, when looking "within-segment", we are back in an analogous situation to Equation (4) in which we ignored the likelihood contribution of how marketing interventions are assigned to units. Equation (9) makes this explicit by writing the conditional density of $\mathbf{x}$ as an r-specific density, $f_{x|\tilde{d}}^{(r)}(.)$ that does not depend on z, and therefore not on $\Omega^{(r)}$. Hence, when estimating this model within segment r, we can ignore the part of the likelihood corresponding to $f_{x|\tilde{d}}^{(r)}(.)$. Since we do not need to know $f_{x|\tilde{d}}^{(r)}(.)$ for inference, we have essentially solved the first problem outlined above.

Secondly, note we estimate the model parameters conditional on being in segment r. All parameters are segment-specific, and we do not do any pooling across segments. To estimate parameters conditional on being in segment r, information on why an observation is in segment r is not required.[8]

[7] A previous working version of the paper incorrectly noted that the sampling of consumers within a segment was purely random. Conversations with MGM's management subsequently revealed that ad-hoc departures from this sampling scheme along the lines described above were possible in some campaigns. We believe this is a small number. The concerns over these are minimized to the extent that we condition on flexible functions of past spending and visitation and its interactions in the model for y. The monte carlo simulations reported in the next section assess sensitivity to this issue.

[8] As an analogy, consider a selection model in which agents self-select into an option, and suppose product usage is observed conditional on selection. To estimate the tastes of the specific sub-population of agents who self-selected into

This means the marginal density of z, which determines segment membership by specifying the probability that $z \in (\underline{z}_r, \overline{z}_r)$, is not part of the segment-specific likelihood (9). Thus, we do not need to know this term for inference of $\Omega^{(r)}$, solving the second problem outlined above.

To summarize, our method for addressing endogeneity essentially divides observations into R non-overlapping segments in a first step, and then estimates separate models for each sub-segment. The key is that the segments are based on thresholds that map to the targeting rule. The reason we are able to do this is we observe the variables on which targeting is based on. More generally, behavioral targeting by firms results in a complicated selection on unobservables problem in estimation for academics who wish to utilize that data for analysis. Observing the variables on which targeting is based on converts the problem of selection on unobservables to that of selection on observables, which facilities controls for nonrandom selection. In practice, we expect this method to work well because it is simple to implement and exploits the internal information of firms which is usually available in such settings. The policy of assigning marketing on the basis of bins of summaries of historical behavior (like Recency-Frequency-Monetary value or other metrics) is common in industry. Hence, we expect our method to have applicability in a variety of contexts. Another advantage is that the informational demands of the method are lower than ones that require knowledge of the full targeting rules employed by firms. The researcher needs to observe the metric (and cutoffs) on which assignment is based. However, since the likelihood of the score can be ignored in estimation, researchers need not know exactly what factors determined the historical evolution of these metrics in the firm's database.

A caveat to this analysis in our setting is we have some uncertainty as to whether the firm exactly adhered to the cutoffs that we were told were used to bin z_{it}. It is possible a slightly different set of cutoffs than we were given were used in the earlier part of our data, when priorities, personnel and systems may have been different. Though our understanding is the extent of change in the cutoffs over the two year period covered by our data is low, we are unable to verify this exactly. Because of this aspect, our method should be seen as approximate. The method will be exact in other situations where researchers know the bin-cutoffs more precisely.

Parameter Interpretation and Prediction To close this section, we discuss parameter interpretation and prediction under our method. An important feature is the procedure above returns z-specific as opposed to individual-specific parameters. In effect, what we are doing is characterizing an individual's type by his z and $\tilde{d}$ and obtaining type-specific parameters. Thus, we make z and $\tilde{d}$ the relevant dimension of heterogeneity. In essence, this is a model of time-varying heterogeneity, in which the variation over time in an individual's parameters is projected onto changes in the relevant summary of his historical behavior, z. A restriction of the method is that the variable onto which heterogeneity is projected has to be the variable the firm uses for its segmentation. Thus, the richer the range of the firm's segmentation policy, the richer the range of heterogeneity the researcher is able to accommodate. If the firm uses a coarse segmentation scheme, the researcher is also able to explore only a similarly coarse range of heterogeneity without making additional assumptions.

A remaining task is to explain how to predict response once the segment-specific estimates are obtained. To predict the response of a given individual i, we would need to know his current average *Theo* (z_{it}) and demographics ($\tilde{d}_i$). The values of these two sets of variables together determine which

an option, one needs to know only usage data of agents conditional on choosing the option; a model of why an agent self-selected into that option is not required for this exercise.

bucket r the consumer currently belongs to. Then, we predict his response using the parameters estimated for segment r, $\hat{\Omega}^{(r)}$. This prediction then serves as an input into a decision-support system or optimizer that allocates a given set of promotions to the set of available consumers in a campaign to obtain the most favorable predicted response at lowest cost.

The reader should be aware of two data requirements of the method, both deriving from the fact that we do not pool across segments. First, there is loss in efficiency from not pooling across segments. This is likely to be a concern in sparse data applications, but may not be practically important in database applications in which there are a large number of observations within each segment. Secondly, pooling across segments has the advantage that one can infer response to the set of marketing offers available to the firm by combining data across all consumers across all segments. Then, information on how other consumers reacted to a marketing offer can be "borrowed" by the statistical model to infer how a focal consumer would react to that offer. This enables a response prediction to be made even if the focal consumer is not observed to be assigned that offer in the data. This advantage is reduced in segment-by-segment estimation. Each segment's response parameters have to be inferred from the behavior of consumers in that segment alone, so any pooling is only across observations within the same segment. An implication is that in order to predict how a segment would react to a given marketing offer, we would need to observe assignment of that offer to at least some consumers in that segment in the data. This is likely to happen with large firms with large prospect databases and/or with large segment sizes. Thus, sparse data settings with fine segment definitions with little marketing variation within segment are unsuitable for application of our method. More generally, when choosing a model for heterogeneity, a researcher always faces the usual tradeoff of utilizing the "within" vs. the "across" variation across units. The within-segment analysis is more credible (we infer parameters only from the behavior of similar units within a segment), however this within variation may be thin. The across-segment analysis is less credible (we have to pool information across very different units), but this across variation may be rich. What method to use is dependent on the analysts' assessment of the tradeoffs between these two considerations. This assessment applies to the decision to use our method as well.

Finally, it is important to note that our method needs to be extended if the goal is to infer long-run response rather than to facilitate short-run campaign optimization. In the long-run, consumers who are targeted promotions as part of a campaign may visit or spend more, causing their future *Theo* to change. This in turn induces a movement into a new *Theo* bucket with new response profiles. To incorporate this, the current set-up will therefore need to be augmented with a model of how *Theo* evolves in response to promotions. Then, Equation (7) that specifies the process for z will have to be fully specified. We avoided modeling this due to model complexity and informational constraints, as our goal was to present a method that does inference without needing to specify the full likelihood. But doing this may become important if the goals become more ambitious and assessments of long-run effects are required.

Intuition for why the Estimator Works We close this section with a simple example that illustrates the intuition for how the procedure works in the context of a linear model. The key to understanding the procedure is to note that we solved the endogeneity problem by holding constant the value of the variable that targeting is based on (z), and then estimating separate models at each of these fixed values. We now explain why this is helpful.

For simplicity, assume a cross-sectional model, let i denote individual, and suppose outcomes y_i, marketing x_i and score z_i are generated in the following way,

$$y_i = \alpha_i + \beta_i x_i + \epsilon_i \qquad \text{and} \qquad x_i = g\left(z\left(y_{i0}, \eta_i\right), \nu_i\right)$$
$$\mathbb{E}\left[\epsilon_i \nu_i\right] = 0$$
$$\mathbb{E}\left[\epsilon_i \eta_i\right] \neq 0$$

where $(\epsilon_i, \nu_i, \eta_i)$ are unobservables to the econometrician that drive y, x and z respectively. Here x causes y; but x is set as an unknown function of z, which in and of itself is a function of i's historical action, y_{i0}. There are some sources of exogenous movement in x represented by ν which are uncorrelated with factors driving y. There are some unobservables driving z; however, those unobservables are correlated with factors driving y (ϵ). This setup is a stylized linear analogue to our model. z is analogous to *Theo*, which is a function of past outcomes. The unmodeled expenditure splits across entertainment options are analogous to η, which drive z and are also correlated with factors driving y. In this set-up, x is endogenous due to z's history dependence, *and* because z is codetermined with y-unobservables. To see how our procedure works, suppose z takes two values, $(0, 1)$. Consider estimation for a fixed value z,

- When $z_i = 0$, $x_i = g_0\left(\nu_i\right)$, estimate $y_i = \alpha_{(0)} + \beta_{(0)} x_i + \epsilon_i$
- When $z_i = 1$, $x_i = g_1\left(\nu_i\right)$, estimate $y_i = \alpha_{(1)} + \beta_{(1)} x_i + \epsilon_i$

In both situations, for a fixed value of z, we use "good variation" in x for identification (here due to ν_i which is uncorrelated with ϵ_i), delivering consistent estimates of z-specific parameters $\{\alpha_{(0)}, \beta_{(0)}, \alpha_{(1)}, \beta_{(1)}\}$. Exact knowledge of $g(.)$ is not required. We would interpret $\beta_{(0)}, \beta_{(1)}$ as the effect of x for the subset of consumers who respectively have $z = 0$ and $z = 1$, i.e., we have projected down to z as the relevant dimension of heterogeneity. This is the strategy used in the paper. While our approach has general application, we expect our proposed strategy to be especially relevant to behavioral targeting situations in Marketing, where marketing interventions are allocated at least partially on the basis of customer history.

5 Monte-Carlo Study

We now present a monte carlo study that investigates the performance of our estimator, and assesses its sensitivity to inexact knowledge by the researcher of bin cutoffs, and to potential nonrandom allocation of marketing interventions to units within bins as reflected in our empirical application. For transparency in interpretation, we use the stylized linear setup presented above, though we modify the setup slightly for ease of illustration of some of the econometric forces driving the results. We start by assuming the true data generating process is as follows:

$$y_i = \alpha + \beta x_i + \epsilon_i \tag{10}$$

We are interested in the recovery of the parameter β. We assume that marketing to unit i, x_i, is allocated on the basis of the bin $k = 1, .., K$ that his score, z_i falls into, as:

$$x_i = \theta_0 + \sum_{k=1}^{K} \theta_k \mathbf{I}\left(z_i \in \mathcal{Z}_k\right) + v_i \tag{11}$$

We model the process generating z_i as,

$$z_i = \omega\eta_i + (1-\omega)\, y_{i0} \tag{12}$$

and assume that,

$$\epsilon_i = \kappa\left(\eta_i; \boldsymbol{\varrho}\right) + \varepsilon_i \tag{13}$$

so that, ϵ_i, the unobservables driving y_i, are directly correlated with the score z_i via ϵ_i's dependence on η_i. In equation (12), ω is a weight such that $\omega \in (0,1)$, and in equation (13), $\kappa\left(\eta_i; \boldsymbol{\varrho}\right)$ are nonlinear basis functions of η_i indexed by parameters $\boldsymbol{\varrho}$. The parameters ω and $\boldsymbol{\varrho}$ control the degree to which z_i (and consequently x_i) is correlated with the unobservables in y_i (larger values of these parameters make x_i more strongly correlated with ϵ_i and accentuates the endogeneity).

We make the following distributional assumptions: $\eta_i \sim \mathbb{U}(0,1)$, $y_{i0} \sim \mathbb{U}(0,1)$, $\upsilon_i \sim \mathbb{N}(0,1)$, $\varepsilon_i \sim \mathbb{N}(0,1)$, all independent of each other. Under these assumptions, z_i lies on the unit line, so we create the bins $\mathcal{Z}_k$ by making K non-overlapping splits of the unit segment. Our simulations vary the parameter vector $\varPsi \equiv \{\alpha, \beta, \boldsymbol{\theta}, \boldsymbol{\varrho}, \omega\}$, the number of bins ($K$) and the number of observations (N). In what follows we will use $N = 10,000$, $K = 3$ and $R = 100$ (number of replications) as the base specification for the purposes of discussion. For most of our simulations, we also assume that $\kappa\left(\eta_i; \boldsymbol{\varrho}\right) = \varrho_1\eta_i + \varrho_2\eta_i^2$.

Performance Metrics

To assess the performance of our estimator, we simulate data from the model, and estimate the below specification separately for each bin k,

$$y_i^{(k)} = \alpha^{(k)} + \beta^{(k)} x_i^{(k)} + \epsilon_i^{(k)} \tag{14}$$

We then construct a pooled estimator, $\widehat{\beta} = \frac{1}{K}\sum_{k=1}^{K}\widehat{\beta}^{(k)}$, and compute the percentage error as,

$$\Delta = \frac{\widehat{\beta} - \beta_{\text{truth}}}{\beta_{\text{truth}}} \tag{15}$$

Under the above model, any of the K estimators $\widehat{\beta}^{(k)}$ should also be unbiased for β_{truth}; so, we also define an analogous percentage error,

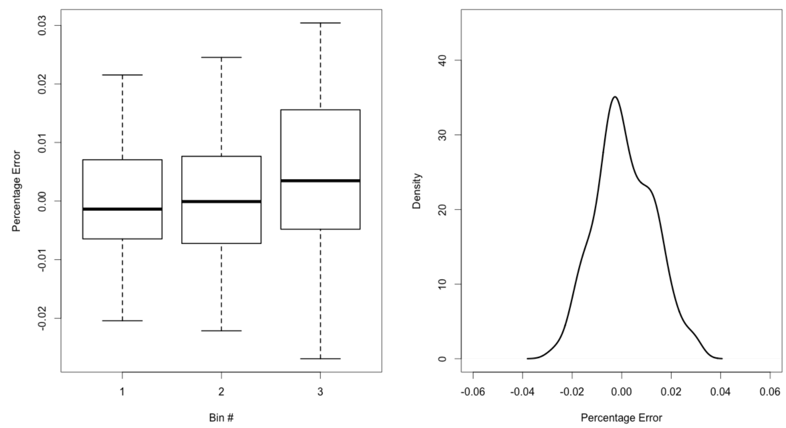
$$\Delta^{(k)} = \frac{\widehat{\beta}^{(k)} - \beta_{\text{truth}}}{\beta_{\text{truth}}} \tag{16}$$

and assess performance on the basis of Δ and $\Delta^{(k)}, k = 1, ..K$.

Results: Base Case

We first demonstrate that our approach recovers β with zero average bias across many replications. In each replication, a true $\varPsi$ vector is drawn randomly, and and a corresponding dataset of $N = 10,000$ units is generated by simulating from Equations (10) to (13) under the assumption that there are $K = 3$ bins. The simulated dataset is then used to estimate the parameters. In this scenario, the researcher is assumed to know the true number of bins and the cutoffs that were used to generate the data, so Equation (14) is estimated separately for each of the K bins in the simulated dataset. Figure (5) plots the implied Δ and $\Delta^{(k)}, k = 1, 2, 3$, across 100 such replications. As is evident from the plot, the bias is small and the recovery of the β parameter is good even if we only took any single bin as our estimate. In simulations not reported, we found that increasing the number of bins $K = \{10, 20, 50\}$ helps improve performance even more, and bias reduces.

Figure 5: Bias in Recovery of β across Replications $\{N = 10,000, R = 100\}$



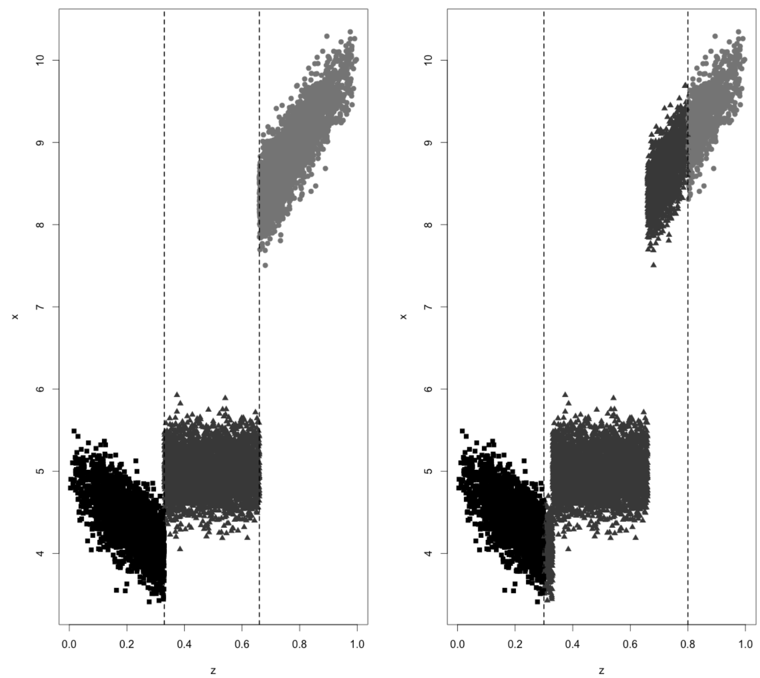
Results: Bin Misspecification

We now explore the estimator's sensitivity to two kinds of possible misspecification: (a) when the researcher incorrectly specifies the cutoffs used in the construction of the bins (i.e., $\mathcal{Z}_k$); and, (b) when x_i is potentially non-randomly allocated within the $z-$bins as a function of past history. To illustrate this situation, we set $K = 3$, and modify the true data generating process for x_i so that it depends directly on z_i and y_{i0} even within the z-bins:

$$x_i = \theta_0 + \theta_1 z_i I(z_i \in \mathcal{Z}_1) + \theta_2 I(z_i \in \mathcal{Z}_2) + \theta_3 z_i I(z_i \in \mathcal{Z}_3) + \delta y_{i0} + v_i \tag{17}$$

Notice that in Equation (17), we do not allow for a direct dependence between x and z in the second bin. We do this so as to illustrate visually (below) the effects of cutoff misspecification; allowing the direct dependence in all bins does not alter the qualitative nature of the results below in any substantive fashion. We simulate $N = 10,000$ units from Equation (17) along with (10), (12) and (13) as before and plot the simulated data in Figure (6). The left panel of Figure (6) shows x against z when the researcher knows the true cutoffs used to generate the data, and the right panel shows x against z when the researcher misspecified the cutoffs. Looking at the left panel, we see that x is strongly dependent on z in bins one and three, but there is little dependence in the second bin. Looking at the right panel, we see that misspecification of the cutoffs however induces a non-randomization in the x variable even in the second bin as seen by the correlation between x and z within the (misspecified) middle bin constructed by the researcher. This, then, is the main effect of bin misspecification $-$ it induces a problematic within-bin dependence between the allocation of marketing and the score, even when such within-dependence does not exist in the true data generating process.

Figure 6: Marketing (x) as a Function of Score (z) with Nonrandom Within-bin Allocation and Misspecified Cutoffs



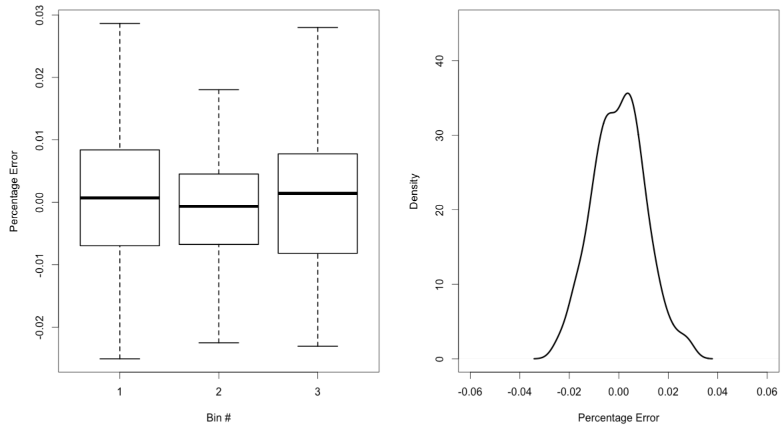
As we noted in the estimation section, we handle this in practice by including z directly in each bin-level model. In particular, now we estimate the following specification that controls for z within each z-bin:

$$y_i^{(k)} = \alpha^{(k)} + \beta^{(k)} x_i^{(k)} + \gamma^{(k)} z_i + \epsilon_i^{(k)}$$

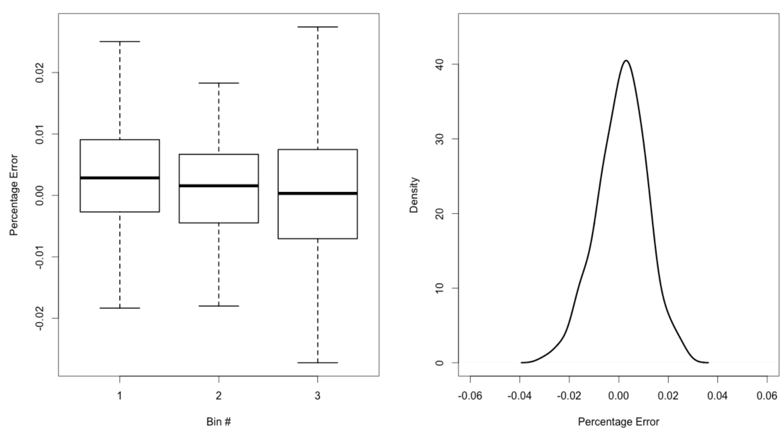
Figure (7) presents the analogous performance plots for this situation across 100 replications. Sub-figure (7a) presents the case with nonrandom within-bin allocation and correctly specified cutoffs; while sub-figure (7b) presents the case with nonrandom within-bin allocation and misspecified cutoffs. We see that the performance of the estimator is quite good in both situations. The intuition for this is that even with the non-randomness and the misspecification of the bins, the variation in x within any bin is driven more by the noise (v) than the confounding variable z. As such parameter recovery remains robust. In simulations not reported here, we found that the performance of the estimator remains good even in cases where the direct dependence between x and z is highly nonlinear and unknown to the researcher, as long as we include a fairly flexible function of $z$directly into the model for y. Finally, we found the estimator can perform poorly in pathological situations where there is little or no independent variation in x within a bin, and where the lack of knowledge of the true cutoffs is so large that the researcher's misspecification bias is overwhelming.

Figure 7: Performance of Estimator with Nonrandom Within-bin Allocation and Misspecified Cutoffs

(a) Nonrandom Within-bin Allocation and Correctly Specified Cutoffs



(b) Nonrandom Within-bin Allocation and Incorrectly Specified Cutoffs



6 Data and Model Operationalization

Prior to operationalizing the model, we spent a significant amount of time and effort on cleaning and scrubbing data to produce a data set fit for estimation. Much of the effort was spent on three aspects, namely, (1) collating different sources of information from disparate units within the company into one central repository (e.g., collating the promotions targeted by the different properties in a particular month together to construct the complete set of promotion options available to every consumer in every month in the data); (2) matching the different sources of information based on unique identifiers (e.g., matching consumer id-s in the transaction database to consumer id-s in the marketing databases of corporate and property-specific departments); (3) cleaning the data to eliminate database coding errors, unreasonable entries and/or missing information.[9]

Data Descriptives

As mentioned above, the data consist of individual-level transactions of a random sample of about 1M consumers picked from MGM's prospect database. Very high-value consumers, who are typically assigned individual hosts and are marketed to separately are not included in this project. Of the consumers in the sample, some are targeted marketing offers, some not. Visitation and transactions of all consumers are observed, so are details of all the offers mailed out and redeemed. Most mass-volume consumers offers are targeted via email or direct-mail. Consumer exposure to print, online and billboard advertising and other media are not included in the data. Hence, some effects of marketing are not captured in our results. To the extent possible, we believe we have captured almost all targeted promotions available to the consumers that are specific to MGM. We believe we have also captured most of the transactions that occur during a consumer visit. Some transaction information is missing if the consumer uses cash or if he does not have a loyalty card number from MGM. But we believe this proportion is small. Transaction information at other casinos outside the MGM family and competitive promotions are not tracked. However, this limitation is shared by all firms in the industry and is a constraint the analytics solution needs to take as given. Developing a database with a 360 degree view of consumer behavior across competing casinos will be an important step forward for the industry as a whole in order to capture competition and "share-of-wallet" better.

Model Operationalization and Details of Variables

Below, we briefly discuss some details of how we operationalized the models in the context of our application.

Segments: We divided observations into $R = 50+$ segments prior to estimation (we do not reveal the exact value of R due to business confidentially concerns).. These segments were based on bins of *Theo*, consumer distance from the casino (Local, Regional, National or International), the number of past trips made by the customer prior to the beginning of the data, and whether the consumer primarily played at slots or tables.[10]

Demographics (d_i): Within each segment, a rich set of demographics (corresponding to d_i) are included in all the estimated models, including age, MGM-specific tier, tenure on books, whether player has a host at MGM (if pertinent), favorite game. In addition, we map in census-level zip code demographic information for each individual into the data including mean household income, disposable income, and mean household expenditure on airfare, entertainment, food & beverage, and lodging.

Past History $\{g(.)$ and $h(.)\}$: To operationalize the functions $g(.)$ and $h(.)$ capturing the history of past play, we include several metrics of past history including average bet, coin-in to point

[9]The 11 MGM properties in Las Vegas which are spanned by the data are: *Aria*, *Bellagio*, *Circus-Circus*, *Excalibur*, *Luxor*, *Mandalay Bay*, *MGM Grand*, *Mirage*, *Monte Carlo, New York-New York and Railroad Pass.*

[10]Tables are more subject to pit-boss effects than slots; hence, incorporating this distinction helps tap into the underlying consumer type better.

ratios, jackpot incidence, number of games played, number of sessions played and time spent. We also include rich functions of past *Theo* and Actual Win including *Theo* and Actual Win at tables and slots separately, in residential casinos versus non-residential casinos, in luxury versus non-luxury properties (based on MGM's definitions of property classifications); *Theo* and Actual Win arising from free play-based play, and other metrics of dollars of incented vs. non-incented play.

Marketing Offers: We include variables measuring the entire range of marketing offers from the company at both the corporate (valid at all properties) and the property-level. These offers are extensive and include:

1. *Room* metrics like room type, room discount, number of comp nights, whether comp is midweek or weekend.

2. *Entertainment, sports and facility* offer metrics like club pool offer, entertainment type, indicator for entertainment offer, ticket price discount, indicator for facility offers, indicator for sports offers, sports offer amounts, sports ticket price discounts, indicator for golf offer.

3. *Casino Event Information* metrics like indicator for inclusion in the casino event prize pool, the prize pool format, indicator for grand prize inclusion, grand prize format, prize value offered, cost of event for which offer is made, buy-in amount, points to entry if offered, tier credits to entry if offered.

4. *Special Event* metrics like indicators for special event, tier upgrade offers, tier credits offered, offers of points that count toward higher tiers in the MGM loyalty program, comps linked to points, point multiplier offers, and multipliers on points that count toward higher tiers (offered on visits that overlap with birthdays).

5. *Retail and spa* offer metrics like indicator for a retail offer, retail offer amount, indicator for spa offer, and spa service amount.

6. *Air and limo* offer metrics like indicator for an airline offer, air package amount, indicator for limo offer, indicator for VIP check-in flag.

7. *Free-play and Promo-chip* offer metrics like free-play offer amount and promo-chip offer amount.

8. *Resort Credit* metrics like resort credit type and resort credit amount.

9. *F&B* metrics like F&B offer and F&B offer amount.

10. *Other* metrics like whether the customer started off his first visit as a result of a database offer, and net reinvestment amount on the consumer.

Table (1) shows the R segment definitions, as well as the proportion of consumers and the number of trips observed in each bin. For confidentiality reasons, an undisclosed, random number of bins are omitted from the table (so the percentages will not add up to 1). Overall, in several of our models, we include over 200+ variables. Estimation of all models described above was programmed in the SAS statistical package.

Table 1: Segment Definitions used in the Analysis

Segment Num.	Segment Bin	Proportion of Consumers	Num. of Trips
1	2+ Trips Local 0-549 Slot	2.04%	206,812
2	2+ Trips Local 0-549 Table	0.25%	24,159
3	2+ Trips Local 0-549 Both	0.13%	13,598
4	2+ Trips Local 550-899 Slot	0.18%	32,450
5	2+ Trips Local 550-1999 Table	0.09%	14,198
6	2+ Trips Local 550-4499 Both	0.07%	12,052
7	2+ Trips Local 900-1999 Slot	0.16%	51,961
8	2+ Trips Local 2000-4499 Slot	0.14%	49,204
9	2+ Trips Local 2000-4499 Table	0.04%	8,981
10	2+ Trips Local 4500-9999 Slot	0.10%	45,146
11	2+ Trips Local 4500-7999 Table	0.02%	6,342
12	2+ Trips Local 4500+ Both	0.02%	9,183
13	2+ Trips Local 8000+ Table	0.02%	18,802
14	2+ Trips Local 10000+ Slot	0.03%	80,961
15	2+ Trips Regional 0-549 Slot	9.81%	496,924
16	2+ Trips Regional 0-549 Table	2.08%	110,485
17	2+ Trips Regional 0-549 Both	1.09%	61,125
18	2+ Trips Regional 550-899 Slot	0.86%	49,538
19	2+ Trips Regional 550-899 Table	0.28%	16,193
20	2+ Trips Regional 550-4499 Both	0.50%	33,632
21	2+ Trips Regional 900-1999 Slot	0.84%	65,138
22	2+ Trips Regional 900-1999 Table	0.28%	21,520
23	2+ Trips Regional 2000-2999 Slot	0.37%	26,017
24	2+ Trips Regional 2000-4499 Table	0.21%	15,659
25	2+ Trips Regional 3000-4499 Slot	0.27%	21,259
26	2+ Trips Regional 4500-5999 Slot	0.16%	12,221
27	2+ Trips Regional 4500-7999 Table	0.10%	7,601
28	2+ Trips Regional 4500+ Both	0.07%	7,629
29	2+ Trips Regional 6000-9999 Slot	0.21%	16,425
30	2+ Trips Regional 8000+ Table	0.09%	17,393
31	2+ Trips Regional 10000+ Slot	0.12%	33,620
32	2+ Trips National 0-549 Slot	25.44%	1,168,783
33	2+ Trips National 0-549 Table	5.00%	228,766
34	2+ Trips National 0-549 Both	2.36%	112,596
35	2+ Trips National 550-899 Slot	2.29%	111,443
36	2+ Trips National 550-899 Table	0.67%	32,319
37	2+ Trips National 550-899 Both	0.54%	24,587

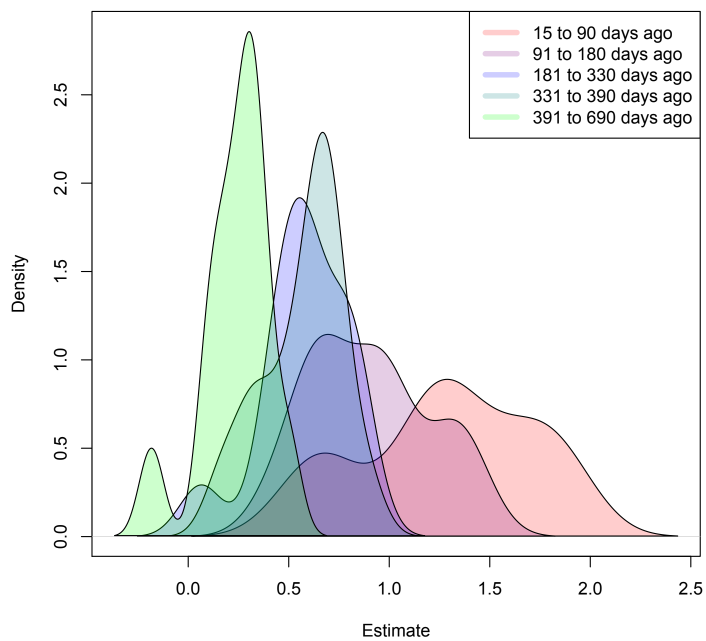
38	2+ Trips National 900-1999 Slot	2.47%	146,007
39	2+ Trips National 900-1999 Table	0.75%	44,398
40	2+ Trips National 900-1999 Both	0.44%	25,752
41	2+ Trips National 2000-2999 Slot	1.03%	55,870
42	2+ Trips National 2000-2999 Table	0.33%	18,416
43	2+ Trips National 2000-2999 Both	0.17%	9,134
44	2+ Trips National 3000-4499 Slot	0.75%	43,683
45	2+ Trips National 3000-4499 Table	0.25%	15,448
46	2+ Trips National 3000-4499 Both	0.11%	6,792
47	2+ Trips National 4500-5999 Slot	0.43%	24,106
48	2+ Trips National 4500-5999 Table	0.16%	9,228
49	2+ Trips National 4500-7999 Both	0.11%	6,506
50	2+ Trips National 6000-7999 Slot	0.33%	19,547
51	2+ Trips National 6000-7999 Table	0.13%	8,308
52	2+ Trips National 8000-9999 Slot	0.22%	12,417
53	2+ Trips National 8000-9999 Table	0.09%	5,768
54	2+ Trips National 8000+ Both	0.07%	9,522
55	2+ Trips National 10000+ Slot	0.34%	54,986
56	2+ Trips National 10000+ Table	0.22%	46,147
.	.	.	.
R	1 Trip International	0.96%	33,449

Notes: Segments based on Number of Trips observed, Range of *Theo*, Location and Slot/Table preference. Data span Jan 2008 to July 2010.

7 Results

We present a representative set of results for brevity. We present parameter estimates and not marginal effects for business confidentiality reasons. The final implementation involves over 120 separate estimated models of consumer behavior (separated by segment, casino and outcomes), over 180+ variables in each model, and about 20,000 parameters estimated across these models. Figure (8) documents the effect of the time since the last trip on visit propensity. We operationalize the effect of time since the last trip in the various models by categorizing it into discrete buckets, and including a dummy variable for each time-interval bucket. To summarize a large number of estimates in a meaningful way, we present the distribution *across* the R segments of each such dummy variable as estimated from the data. Figure (8) presents these distributions. Looking at Figure (8), we see strong evidence of duration dependence in the data. The hazard of visitation is in general declining in the time since the last visit: those that visited 15-90 days (pink distribution) are on average roughly 6 more likely to visit than those who have visited more than 391-690 ago (lime green distribution). This may also reflect within segment heterogeneity in that the first bucket comprise consumers with high utility

Figure 8: Effect of Time Since Last Trip on Visit Propensity



from gambling and visitation, while the second reflect those with lower value (or costs) from visits. These duration effects allow the model to link a customers' visitation behavior over time in assessing his relative value to the casino. For business confidentiality reasons, we cannot report how these numbers or the ones below translate into visit propensities at each of the individual properties or into profitability or revenue.

Figure (9) presents the effect of the money won on the previous trip on current visitation propensity. We include the money won on the previous trip as a continuous variable in all models. Figure (9) plots a histogram of the coefficient on this variable across the R segments. Figure (9) documents interesting heterogeneity of past winnings on current visitation: there is a bi-modal distribution of effects, with a smaller segment for whom past winnings seem to matter significantly in driving future visits. This group may form a viable segment for targeting free-play promotions, for instance. Figure (10) documents the effect of the distance from the consumer's residence to Las Vegas on visit propensity. We operationalize this as a continuous variable that varies across consumers included in each segment (note that even though we create segments based on distance, we still have variation in distance across consumers within each segment). Figure (10) plots a histogram of the coefficient on this variable across the the R segments. Interestingly, we find the effect of distance is not uniform: for some segments, especially those within the "Regional" and "Local" distance segments, living further away increases visit propensity, perhaps capturing satiation with gambling opportunities or the characteristics of suburban gamblers.

We now discuss some results on the effect of targeted marketing. As a representative example, we plot the effect of providing a free room on visitation. We estimate a separate effect of a free

Figure 9: Effect of Money Won in Last Trip on Visit Propensity

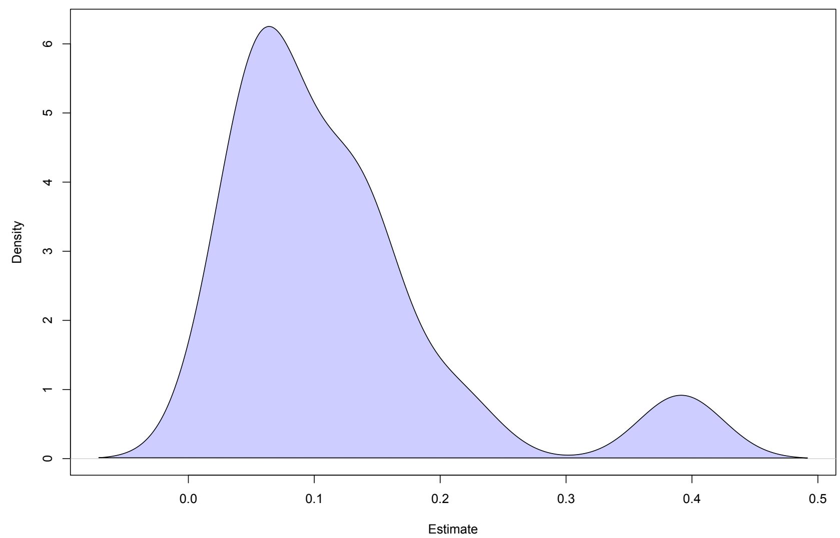


Figure 10: Effect of Distance to Vegas on Visit Propensity

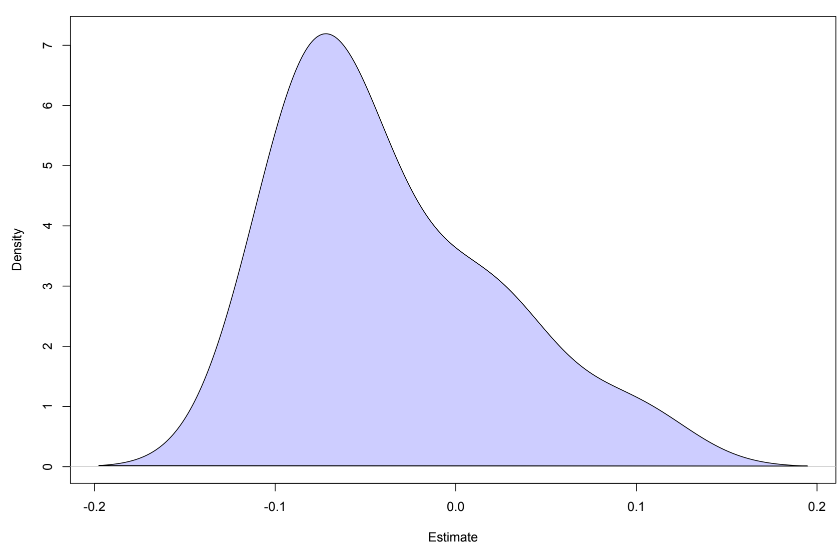
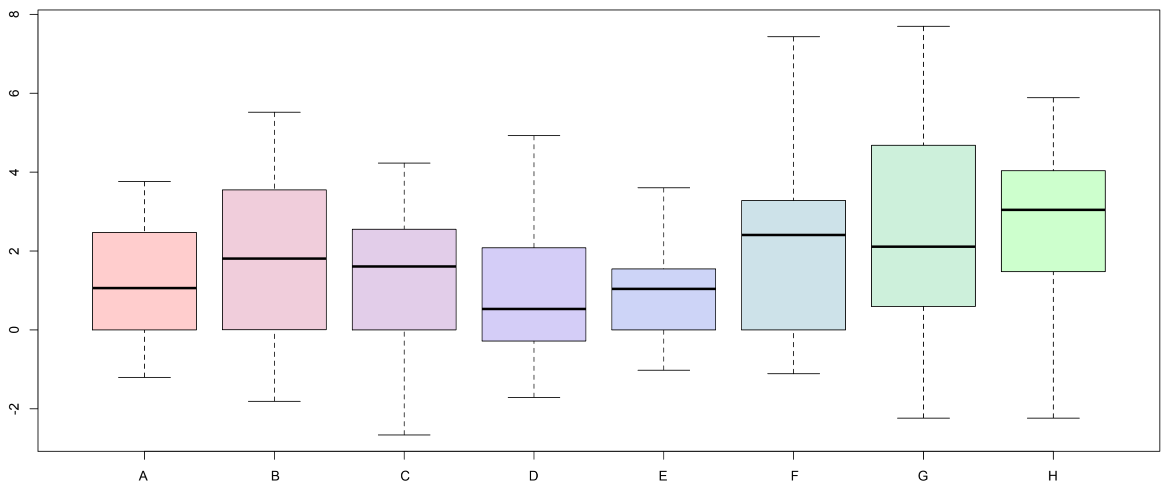


Figure 11: Effect of Free Room (Relative to Property X)



room promotion at each casino and for each of the R segments. We operationalize these effects by interacting a free-room dummy, with a dummy for which casino the free-room can be availed at, and including these interaction variables in the model of visitation for each of the R segments. In Figure (11) we plot the estimated effect of providing a free room at a given casino relative to providing a room at one of the casino properties, called property X. Each box-plot presents the distribution of that casino's effect relative to property X plotted across the R segments. For example, the box-plot on the extreme left of Figure (11) named "A" shows the distribution across the R segments of the effect of providing a free room at property A relative to that at property X. Interestingly, the effects are all positive, implying that providing a free room at each of the listed casinos has a higher effect on visitation relative to providing one at property X, suggesting that free-room provision at property X produces little marginal visitation relative to the others. By allowing for heterogeneous property-specific promotions in this manner, the model helps assess the property-promotion-customer match better, so as to result in better optimization of promotions across customers and properties in a subsequent stage.

Finally, we also present plots of the effect of customer characteristics on spending conditional on visit. Figures (12a) and (12b) present the effect of customer age and gender on spending. To operationalize customer age and gender in our spending model, we create dummy variables for various age buckets, interact these with gender (Male/Female dummy), and include these interacted dummy variables in models of spending for each of the R segments. This produces flexible specifications of demographic effects. In the left panel in Figure (12a) we plot the effect of customer age on spending relative to that of the "less than 25 years" bucket for Males. Each box-plot presents the distribution for males across the R segments of being in that age bucket relative to customers who are less than 25 years old. For example, the box-plot on the extreme left of Figure (12a) shows the distribution across the R segments of the effect of being a male aged 25-35 relative to a male aged <25 yrs. on spending propensity. Figure (12b) shows the analogous plot for females. Interestingly, we see little systematic differences in spending all things held equal, across various age tiers for males. However, the distribution is inverted U-shaped for females: women aged 25-35 are significantly less likely to spend compared to those below 25 years; older women are more likely to spend; while spending drops to the base level for the oldest bucket. These demographic differences in spending captured by the model are utilized in improving the match between promotions and customers in the subsequent optimization steps.

We presented only a flavor of the results given space and confidentiality considerations. The main point is that at the end of this process, we have at our disposal a set of empirical models that predict a person, property and trip specific promotional-lift for each available promotion or promotion bundle. These predictions form inputs into the second module of the analytics solution, as discussed below.

8 Optimization

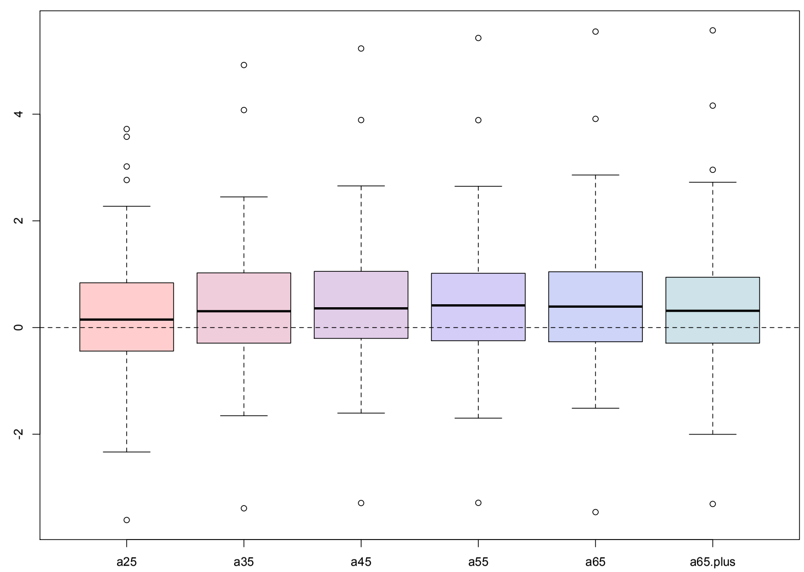
The second module involves an optimization platform that searches within a specified promotion set to find the right promotion bundle for each consumer in the database. The optimization is implemented at the campaign-level and is operationalized in the following manner. First, managers decide on the goal of an advertising campaign (e.g., drive visitation at a new property, increase play at slots etc.). Based on the goals of the campaign, the firm decides on a set of component promotional options that could potentially be offered to the customer base (e.g., a given level of discount at the new property or a given level of free-play credits). Taking these components as given, each customer is scored on each possible component or bundle of component promotions. The scoring is based on the models of visitation and spending outlined above. The score captures the expected profitability of the customer if given the promotion option, by computing expected spending (unconditional on a visit), and subtracting out the expected cost to the firm of offering that promotion bundle to the customer. At the end of the scoring step, we have for each customer a recommended promotion bundle that yields the highest expected profits. In some situations, the optimization may be constrained by additional decision rules desired by management. For example, management may impose a decision rule of not offering more than two promotional components in a bundle for consumers of a particular type; or alternatively, impose a minimum margin the promotion has to meet in order for a customer to be eligible. Figure (13) depicts the process developed for optimization.

The main requirement for an optimization package that implements the above methodology is the ability to scale rapidly to scoring large numbers of consumers. For instance, a specific campaign we consider in the next section involves scoring about 1.5M consumers on close for 50+ promotional options. The optimization package should also integrate well with the statistical models presented above, and provide managers to ability to add constraints to the optimization in a user-friendly manner. We implemented these on Teradata©, a commercial database platform that is developed for Big Data analytics applications.

On the optimization front, the main gain is the ability to customize promotions to each individual consumer. Prior to our engagement, promotions were assigned at the segment-level. The new system enabled optimizing promotions to the individual consumer-level, facilitating finer micro-targeting. Further, compared to the prior system, the number of bundles that could be considered increased by about 6X, increasing the number of instruments available to improve promotion efficiency. With additional hardware space and time, it is straightforward to scale up the system to accommodate even

Figure 12: Effect of Customer Age on Spending

(a) Males



(b) Females

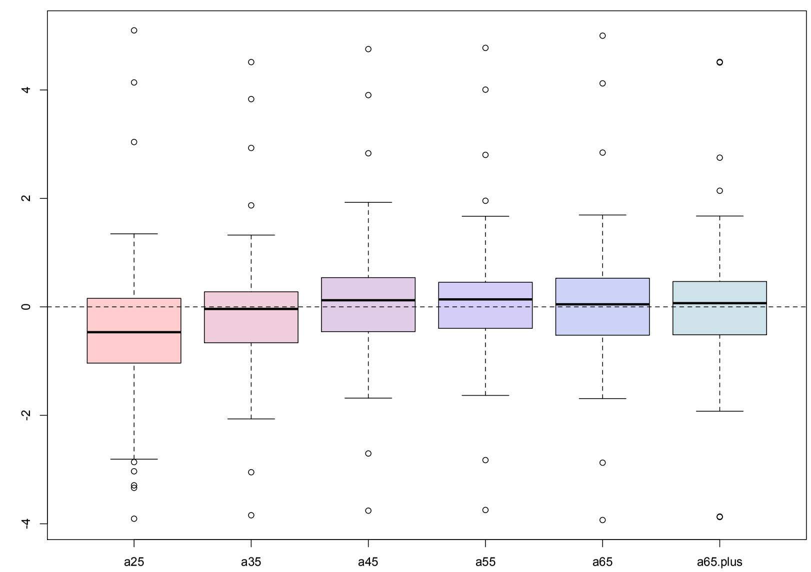
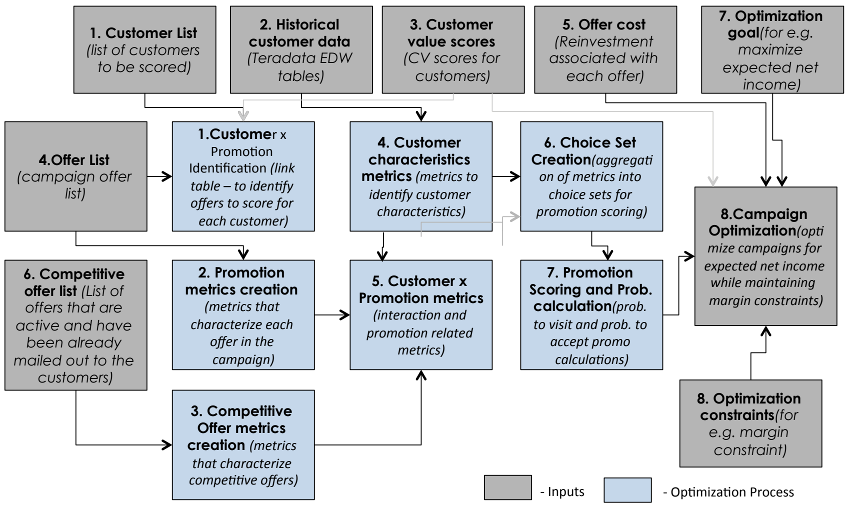


Figure 13: Optimization Process



larger bundle sets.

Finally, dashboard applications were also developed that enabled managers to monitor and dissect company performance on their desktops. Figure (14) provides an example. These dashboards were linked to the underlying statistical model and optimization packages, so as to embed the framework in a user-friendly decision support system.

This completes the discussion of the model framework and development effort.

9 Results from a Randomized Evaluation

The models developed above are assessed as part of a large-scale randomized test at MGM. We implement the test as part of the Summer 2012 corporate campaign. The test evaluates the performance of the model in selecting consumers to whom a set of promotion bundles could be targeted, as well as the ability of the model to match consumers to one of these promotion bundles.

The test involves about 1.5M customers, picked from MGM's prospect database. The goal of the test is to compare the efficacy of the model in promotional targeting relative to the status quo approach used at the firm. The test is implemented as follows. First, a pilot test was conducted in Spring 2012 with a limited number of promotional offerings to assess test design, understand ballpark customer response and to understand logistics of implementation. Based on this, the 1.5M consumers in the prospect data are randomly divided into 3 groups prior to the beginning of summer. Group 1 consumers (30% of the total) are scored on the model we develop. Group 2 (30% of the total) consumers are scored based on the status-quo approach (*Theo* on last visit and demographics). Group 3 consumers (10% of the total) are treated as the control – they do not receive any of the corporate offers. The remaining 30% of consumers are tested on auxiliary aspects that are unrelated

Figure 14: Screenshots of Dashboard Applications for Campaign Management and Monitoring Linked to underlying Empirical Models

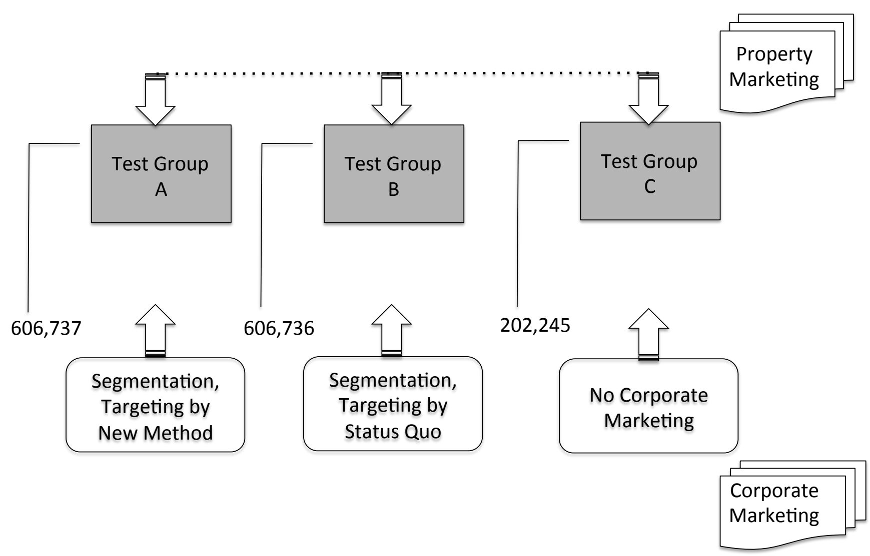
to the model, and are not relevant to this evaluation. Managers then created 75+ promotion bundles which comprise the set of possible promotions that could be offered to consumers in the campaign. Each consumer could be offered only one bundle during the campaign. Once scoring is completed for groups 1 and 2, assignment of these promotion bundles is implemented in the following way. In group 1, each customer is assigned the promotion bundle that provides the highest profitability subject to making a minimum threshold margin. The margin is set outside the model by management. If the expected profitability of the best offer does not meet the margin threshold, no corporate promotion is sent to that consumer. In group 2, consumers are sorted into segments as per the status quo method. Managers then assign a segment-specific promotion bundle to each segment in the group in the same way as they have made those decisions in the past. All corporate offers are then emailed to consumers. The offers are valid for redemption from July 31 to October 31, 2012. All visits to any of the MGM properties along with all transactions involving any of the 1.5M consumers are then tracked during the July 31 − Oct 31 window during which the promotion is active. No other corporate marketing was targeted at these consumers at that time.

During the same period, the three groups are also exposed to *non-corporate* campaigns set independently by the MGM resorts' individual property marketing teams. Thus, those in the control and treatment groups receive other property-specific promotion offers over and above those associated with the corporate test. Hence, the performance of groups 1 and 2 relative to the control should be interpreted as the net effect of the corporate campaign relative to the existing property-specific marketing activity. The promotions for the individual properties are mailed out in the Summer prior to July, and are set independent of the corporate campaign; so they are not jointly allocated or adjust in respond to the corporate level interventions in the test and control groups.[11] These aspects of the test-design derives from organizational considerations within the firm − organizationally, it was not feasible to stop all property-specific promotional activity during the test period. Therefore, we believe the test provides a lower bound on the performance of the new methods within the firm as a whole, because further gains to what is assessed in the test can be had if property-specific campaigns are also coordinated with the corporate campaign. Further, to the extent the new method is better, implementing it at the individual properties as well would be the source of additional gains from the adoption of the new analytics solution. Figure (15) presents the test design pictorially.

Table (2) reports the results. For business confidentiality reasons, all dollar numbers in Table (2) have been scaled by an undisclosed constant; so these should be interpreted as scaled dollars. When we refer to revenues or costs below, we refer to these in scaled dollars, which we call units of "R$" for brevity. Column 1 of Table (2) reports on the results for Group 1, column 2 for Group 2 and the last for the control. The top row of the table reports "adjusted" revenues for each group. These are constructed by summing all gaming and non-gaming *Theo* from agents in a group that visited during the July 31 − Oct 31 window, and subtracting out any Free-play dollars used as well as the dollar value of any MGM-specific reward points redeemed during that period. The company does not count free-play and reward point redemption as sources of real revenue, and considers this as the right metric for assessing policy. Note this makes the adjusted revenues a more conservative metric of the gains

[11] Individual-properties also employ on-the-floor promotions like free-drinks allocated to playing patrons, that may adjust to the consumer play behavior induced by our interventions. While we do not rule this kind of promotion adjustment at the individual-property level, we believe the impact of these on our results is very small, as these kinds of activities account for less than 5% of promotional spending by the properties. The bulk of property specific promotions are mailed out and are pre-determined during the intervention period.

Figure 15: Test Design



Notes: The figure shows the design of the test conducted to evaluate the proposed model. Group A consumers were offered corporate promotions based on the model; Group B based on the status quo method; Group C (Control) consumers were offered no corporate promotions. All groups continued to receive promotions offered by individual properties. The offers are valid for redemption from July 31 to October 31, 2012 and are mailed out in Summer 2012. All visits to any of the MGM properties align with all transactions involving any of the 1.5M consumers in the test are then tracked during the July 31 – Oct 31 window during which the promotion is active.

from a campaign. The next row reports the costs of the campaign across the three groups. These are calculated as the net dollar value of promotions redeemed.[12] Other costs of running the campaign (e.g., printing direct mail) are not included. The costs row for groups 1 and 2 refer to the costs incurred by MGM via redemption of either corporate promotions or property-specific promotions assigned to consumers in that group (unfortunately we are unable to split these out separately by corporate versus property-specific redemptions). The cost entry for the control group refer to the costs incurred by the properties to run the other campaigns they conducted in parallel to the focal corporate promotion.

Looking at Table (2), we see that net adjusted revenues from those treated under the status-quo policy amounted to about R$111.97M, compared to R$114.06M under the new model. Thus, adjusted revenues are higher under the new policy.

We also see net costs are about R$41.42M under the new policy versus R$43.90M under the status-quo. Thus, the new policy makes more money at lower costs.

The upshot of the revenue and cost implications is about R$72.64M profit to the casino under the new policy compared to about R$68.07M under the status-quo. The difference is about R$4.57M for this campaign. Even though we cannot disclose how much this is in real dollars, we are able to disclose a range − in real dollars, this incremental difference is between $1 and $5M incremental profit for the firm.

The comparison to the control group is also informative about the relative profitability of the new method compared to the campaign strategies of the individual properties. Looking at the third column in Table (2), we see that the various other campaigns run concurrently by the individual properties brought in about R$36.8M of revenue from consumers in the control group. Recall that the control is $1/3^{rd}$ the size of groups 1 and 2; so to obtain a relative comparison, we should multiply the dollars in the control by 3. Computing scaled revenues 3×R$36.8M = R$110.4M, we see the new method is superior in terms of revenues to the aggregated impact of the individual property campaigns as well, bringing in about R$3.7M more (R$114.1M for the new policy vs. R$110.4M for the control). Computing costs, the individual properties spent a scaled total of 3×R$14.5M = R$43.5M. The net profit impact is 3×R$22.2M = R$66.6M, which is less than the R$72.6M profit associated with the new policy. Note these comparisons are at the aggregate level, comparing the new method to the sum total of the effect across all 12 properties, and not a comparison of any one property's method.[13]

Computing a Return on Investment per dollar spent, we find the new policy provides a net ROI of about 2.75 compared to 2.53 for the individual properties, and 2.55 for the status-quo approach. Thus, a dollar spent in promotions generates about 20 cents more incremental spending under the new policy compared to the current practice at the firm. As another metric, if the 4 campaigns in one year

[12] These may not map exactly to the true economic cost of the promotion in some cases. For example, the opportunity cost of providing a free room may be lower than the current price of the room if capacity constraints are not binding. Nevertheless, because exact information on these costs are not available, we use the dollar value of the promotion as the measure of costs. These is the metric used at the firm as well.

[13] A related question here is why the redemption costs in groups 1 and 2 are less that the scaled costs of the control group, even though both groups 1 and 2 are also exposed to similar property-level promotions as the control. The reason is that consumers in groups 1 and 2 can choose which promotion − property or corporate − to redeem during their visit to the casino, while consumers in the control can only use property-specific promotions. Even though consumers in group 1 (respectively group 2) are allocated the same property-specific promotions as the control group, they may choose to use the corporate promotion targeted to them during their visit. It could well be that the corporate promotions a consumer utilizes are less expensive to MGM than the property-specific ones offered. An example can help clarify. Suppose a household in group 1 is assigned property-specific promotions in the form of free tickets to an expensive show. If that household is traveling as a family with children, it may prefer a suite upgrade to a free show, though its dollar value is lower. So if a corporate suite upgrade-based promotion is offered, it is more likely to be utilized, inducing lower redemption costs.

each spent the same amount on promotions as the Summer campaign, at these levels of ROI, the firm would make about R$33.6M (or between $10M and $15M in real dollars) in incremental profit from using the new model compared to the status-quo method, or about R$41.7M (or between $14M and $19M in real dollars) in incremental revenues compared to the aggregation of the campaign planning strategies of the individual properties.

The source of the improvement arises from two factors, one due to the improved matching of promotion types to household preferences implied by the model, and two, from the reallocation of promotions from average to marginal consumers (who are more likely to respond to the promotion). Table (2) allows us to informally assess which is the stronger force. If our model simply reallocates the same promotions as before to consumers who are more likely to respond to them, we would have seen that redemption costs went up (or weakly remained the same), and revenues weakly increased. The fact that we see costs go down and revenues went up suggests that better matching plays an important role in the profit improvement in addition to reallocation.

10 Conclusions

Efforts on developing and implementing a comprehensive marketing analytics solution for a real-world company is presented. The framework leverages the richness of the company's data to develop detailed models of consumer behavior for use for optimized targeting. The models feature themes emphasized in the academic marketing science literature, including incorporation of consumer heterogeneity and state-dependence into utility, and the development of new methods for controlling for the endogeneity of the firm's historical targeting rule in estimation. The issues discussed are relevant for other customer-facing firms operating in data-rich environments that wish to improve their promotion targeting and management using modern quantitative methods. The models are then assessed relative to the status-quo using a large-scale field intervention that shows the profits from adopting the new system are substantial. We believe the scale of model development and implementation and the combination of the econometrics with a field intervention in the context of a real-world firm are novel to the marketing science literature.

For academics wishing to port marketing science models from theory to practice, our experience in model building holds a few lessons. First, heterogeneity in consumer behavior is extensive. A large number of segments are required to capture the amount of heterogeneity seen in the data. Even with $R = 50+$ segments, we could detect significant amount of within-segment heterogeneity, some of which we do not model simply on account of practical difficulties, and for ease of model implementation and simplicity in use and exposition. We try to capture much of this by including functions of past behavior into the model (along with demographics).

Second, even with this extensive segmentation, we have a large number of observations in each bucket, a luxury afforded by the Big Data revolution of recent years. Because of this, we found we are able to estimate most of our effects fairly precisely and that issues of sampling error associated with data sparsity, often a significant issue in academic work, are not at the forefront in this setting. Rather, practical significance and not statistical significance becomes the salient issue in model selection. Additionally, the availability of large quantities of data facilitated a "within-segment" or more "local" analysis, as opposed to having to pool across differently targeted units. This will increasingly become possible as we get more unit-specific data collected at scale.

Table 2: Aggregate Performance of Treatment and Control Groups

	New (N=606,736)	**Status-Quo (N=606,737)**	**Control (N=202,235)**
Adjusted Revenues	R$114.06M	R$111.97M	R$36.77M
Costs	R$41.42M	R$43.90M	R$14.53M
Margin	63.68%	60.79%	60.49%
Profit	R$72.64M	R$68.07M	R$22.24M
ΔProfit (A - B)	R$4.57M	–	–
Return on Investment (ROI)	$2.75	$2.55	$2.53

Notes: Group 1 consumers were offered corporate promotions based on the model; Group 2 based on the status quo method; Control consumers were offered no corporate promotions. All groups continued to receive promotions offered by individual properties. Cost of campaign calculated as net value of promotions offered. Other costs of running the campaign are not included. The costs row for groups 1 and 2 refer to the costs incurred by MGM via redemption of either corporate promotions or property-specific promotions assigned to consumers in that group. The cost entry for the control group refer to the costs incurred by the properties to run the other campaigns they conducted in parallel to the focal corporate promotion. ROI is calculated as $\frac{Revenues}{Cost}$.

Third, fitting the data well often requires inclusion of variables numbering in the hundreds. More generally, the advent of database marketing, the integration of marketing with technology, and the proliferation of instruments by which to incentivize and reach consumers imply that a large number of marketing metrics are now tracked at firms. Incorporating these into econometric models while allowing for flexible specifications that include main and interaction effects imply systems with hundreds of variables. Thus, software that can manage the scale of both data and the variable set become key. Also key are statistical methods that scale well. For instance, maximizing a likelihood over a large set of parameters with a large number of observations becomes quickly problematic if the objective functions are not smooth and the likelihood is not concave. Hence, in such situations, models like the logit that have well defined, smooth and concave objective functions, and scale well in variables, become vey attractive. As the number of variables to be considered increases, variable selection algorithms like LASSO will also increasingly become more important in practical applications.

Fourth, we find that thinking structurally about the data generating process is very important in assessing the historical variation in the data and in using it to formulate policy. In our setting, we found that accommodating the firm's historical targeting rule in inference was critical to measuring the right effect of promotions, and for guarding against recommending a significant ramping up of promotions when using the estimated parameters for formulating marketing policy.

Fifth, in many large organizations, getting an analytics project off the ground involves a significant fixed cost associated with data collation and cleaning. It is typical that data are spread across various units within the organization, that some parts of the data are "dirty" or missing, and that some data are available only in unstructured or non-digital form. Thus an academic or consulting company engaging in an analytics effort with the organization should expect to invest a significant amount of upfront time and effort in data cleaning and scrubbing. In our view, this component of the engagement is of critical importance, and reaps large investments because the value of the subsequent modeling is driven to a great degree by the richness and quality of data inputs.

11 References

- American Gaming Association (AGA) (2012), "State of the States: The AGA Survey of Casino Entertainment," `http://www.americangaming.org/industry-resources/research /state-states`, accessed October 17, 2012.

- Ansari, A. and Mela, C. (2003). "E-customization," Journal of Marketing Research, 40(2):131{145.

- Arora, N., Dreze, X., Ghose, A., Hess, J. D., Iyengar, R., Jing, B., Joshi, Y., Kumar, V., Lurie, N., Neslin, S., Sajeesh, S., Su, M., Syam, N., Thomas, J., and Zhang, Z. J. (2008). "Putting one-to-one marketing to work: Personalization, customization, and choice," Marketing Letters, 19(3-4):305{321.

- Bazelon., C, Neels, K. Seth, P. (2012). "Beyond the Casino Floor: Economic Impacts of the Commercial Casino Industry," The Brattle Group, `http://www.americangaming.org/industry -resources/beyond-the-casino-floor`./..6/.54

- Bucklin, Randolph E. and Sunil Gupta (1999), "Commercial Use of UPC Scanner Data: Industry and Academic Perspectives," Marketing Science, 18 (3), 247–73.

- Chiang, J. (1995). "Competing Coupon Promotions and Category Sales," Marketing Science, 14:1, 105-122.

- Cho, S. and J. Rust (2008), "Is Econometrics Useful for Private Policy Making? A Case Study of Replacement Policy at an Auto Rental Company," Journal of Econometrics 145 243257.

- Encyclopedia.com (2012), "Commercial Casinos," `http://www.encyclopedia.com/topic/ Casino.aspx`, Accessed Oct 17, 2012.

- Goldfarb, A. and Tucker, C. E. (2011), "Online Display Advertising: Targeting and Obtrusiveness," Marketing Science, pages 1-16.

- Hartmann, W., Nair, H. and Narayanan, S. (2011). "Identifying Causal Marketing Mix Effects Using a Regression Discontinuity Design," Marketing Science, 30(6), Nov-Dec, pg. 1079-97.

- Lilien, G. John Roberts and Venkatesh Shankar (2013). "Effective Marketing Science Applications: Insights from the ISMS Practice Prize Papers and Projects," Marketing Science Vol. 32, No. 2, March-April, pp. 229-245.

- Little, John D. (1970) "Models and Managers: The concept of a decision calculus." Management Science. 16(8):B-466–B-486.

- Leeflang, Peter S. and Dick R. Wittink (2000), "Building Models for Marketing Decisions: Past, Present and Future," International Journal for Research in Marketing, 17 (2/3), 105–126.

- Lodish, Leonard M. (2001), "Building Marketing Models that Make Money," Interfaces, 31 (3), S45–S55.

- Manchanda, P., P. E. Rossi and P.K. Chintagunta. (2004), "Response Modeling with Non-Random Marketing Mix Variables," Journal of Marketing Research, 41 (November), 467-478.

- Mantrala, M., P.B. Seetharaman, Rajeev Kaul, Srinath Gopalakrishna and Antonie Stam (2006), "Optimal Pricing Strategies for an Automotive Aftermarket Retailer," Journal of Marketing Research, 43, 4, 588-604.
- Misra, Sanjog and Harikesh Nair. (2011). "A Structural Model of Sales-Force Compensation Dynamics: Estimation and Field Implementation," Quantitative Marketing and Economics, 9 (3), September, pg. 211-25.
- Narayanan, Sridhar and Puneet Manchanda (2012). "An Empirical Analysis of Individual Level Casino Gambling Behavior," Quantitative Marketing and Economics, Vol. 10, No. 1, pp. 27-62.
- Roberts, John (2000). "The Intersection of Modeling Potential and Practice," International Journal of Research in Marketing, 17 (2/3), 127–34.
- Rossi, P. E., McCulloch, R. E., and Allenby, G. M. (1996). "The Value of Purchase History Data in Target Marketing," Marketing Science, 15(4):321{340.
- Sahni, N., Zou, D. and P. Chintagunta, (2014), "Effects of Targeted Promotions: Evidence from Field Experiments," working paper, Stanford GSB.
- Simester, Duncan, Peng Sun and John Tsitsiklis (2006), "Dynamic Catalog Mailing Policies," Management Science, 52(5), 683-696.
- Simester, D. Y. Jerey Hu, E. Brynjolfsson, and E. Anderson, (2009). "Dynamics of Retail Advertising: Evidence from a Field Experiment," Economic Inquiry, 47(3):482-499.
- Sinha, Prabhakant and Andris A. Zoltners (2001). "Sales-force Decision Models: Insights from 25 Years of Implementation," Interfaces, 31 (3), S8–S44.
- Van Bruggen, Gerritt H. and Berend Wierenga (2001), "Matching Management Support Systems and Managerial Problem-Solving Modes: The Key to Effective Decision Support," European Management Journal, 19 (3), 228–38.
- WCAI (2014). "Successful Applications of Customer Analytics Conference", `http://wcai.wharton.upenn.edu/conf14/`. (accessed Sept 4, 2014)
- Winer, R. S. (2000). "Comments on Leeflang and Wittink," International Journal for Research in Marketing, 17(2–3):141–145.
- Zantedeschi, D., Feit, E., Bradlow, E. (2014). "Measuring Multi-Channel Advertising Response Using Consumer-Level Data," working paper, Wharton School of Business.