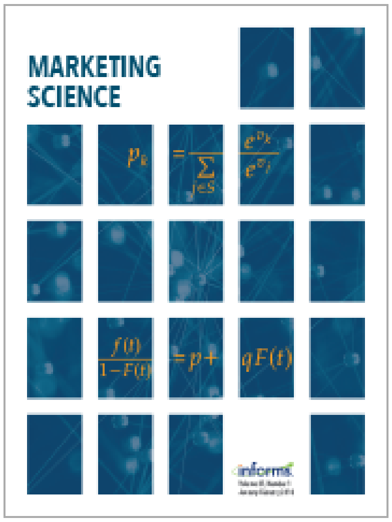


Marketing Science

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

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

Harikesh S. Nair, Sanjog Misra, William J. Hornbuckle IV, Ranjan Mishra, Anand Acharya

informs
http://pubsonline.informs.org/journal/mksc/
MARKETING SCIENCE
Vol. 36, No. 5, September–October 2017, pp. 699–725
ISSN 0732-2399 (print), ISSN 1526-548X (online)

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

Harikesh S. Nair,[a] Sanjog Misra,[b] William J. Hornbuckle IV,[c] Ranjan Mishra,[d] Anand Acharya[d]

[a] Stanford Graduate School of Business, Stanford University, Stanford, California 94305; [b] Chicago Booth School of Business, University of Chicago, Chicago, Illinois 60637; [c] MGM Resorts International, Las Vegas, Nevada 89109; [d] ESS Analysis, Newton, Massachusetts 02466

Contact: harikesh.nair@stanford.edu, http://orcid.org/0000-0003-1697-7767 (HSN); sanjog.misra@chicagobooth.edu (SM); bill_hornbuckle@mgmresorts.com (WJH); ranjanmishra@essanalysis.com (RM); anandacharya@essanalysis.com (AA)

Received: September 6, 2014
Revised: October 27, 2015; December 12, 2016
Accepted: January 13, 2017
Published Online in Articles in Advance: July 31, 2017

https://doi.org/10.1287/mksc.2017.1039

Abstract. Efforts on developing, implementing, and evaluating a marketing analytics framework at a real-world company are described. The framework uses individual-level transaction data to fit empirical models of consumer response to marketing efforts and uses these estimates to optimize segmentation and targeting. The models feature themes emphasized in the academic marketing science literature, including incorporation of consumer heterogeneity and state dependence into choice, and controls for the endogeneity of the firm's historical targeting rule in estimation. To control for the endogeneity, we present an approach that involves conducting estimation separately across fixed partitions of the score variable that targeting is based on, which may be useful in other behavioral targeting settings. The models are customized to facilitate casino operations and are implemented at the MGM Resorts International's group of companies. The framework is evaluated using a randomized trial implemented at MGM involving about 1.5 million consumers. Using the new model produces about \$1 million to \$5 million in incremental profits per campaign, translating to about 20¢ in incremental profit per dollar spent relative to the status quo. At current levels of marketing spending, this implies between \$10 million and \$15 million in incremental annual profit for the firm. The case study underscores the value of using empirically relevant marketing analytics solutions for improving outcomes for firms in real-world settings.

History: Preyas Desai served as the editor-in-chief and Bart Bronnenberg served as associate editor for this article.
Supplemental Material: Data are available at https://doi.org/10.1287/mksc.2017.1039.

Keywords: marketing • promotions • casinos • behavioral targeting • nonrandom targeting • endogeneity • field experiments

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 microtarget 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 document that the use of the data in combination with empirical models of consumer behavior and a large-scale optimization system improves the return on investment (ROI) of marketing effort at the firm and increases the profitability of targeted marketing.

In a randomized field evaluation involving about 1.5 million consumers in the firm's database, we find the new system produces between \$1 million and \$5 million 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 toward marginal consumers for whom the promotion has an incremental impact, and from the improved matching of promotion types to consumer types. Computing a return per dollar spent, we find the new policy provides a net incremental ROI of about 0.20 compared to the status quo policy. Thus, a dollar spent in promotions generates about 20¢ more incremental profit using the model compared to the current practice at the firm. Assuming the same level of marketing spends, this translates to approximately between \$10 million and \$15 million in incremental annual profit from shifting from the status quo to the new model. Taken together, we believe these numbers suggest the new policy is successful.

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, Nevada. 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 the spring and summer of 2012. The project fits into the rising trend usage of marketing analytics in customer facing industries, including the gaming industry. The project is part of MGM's initiatives to use data 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 framework is built on standard empirical models of consumer behavior, which have their genesis in the marketing science literature. While the models are tailored to the casino and gaming setting, aspects of the models that are more general include incorporation 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. 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.

The data reflect historical behavioral targeting by the firm. To use these data for estimation of marketing effects, we develop a strategy to address the endogeneity implied by the nonrandom historical targeting. It involves estimating separate response models for fixed partitions of the historical score variable on which targeting is based. This strategy may be useful in other situations in which researchers work closely with firms and have some knowledge of the firm's past targeting rule. The final implementation involves about 120 separate estimated models of consumer behavior (separated by segment, casino, and outcomes), about 180+ variables in each model, and about 20,000 parameters estimated across these models. We believe the scale of the model implementation and the evaluation of its impact via a large-scale randomized field experiment are novel to the academic marketing literature.

Our study adds to a 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. (2006) on catalog mailing, and Misra and Nair (2011) on sales force compensation. The 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, the paper by Lilien et al. 2013, who review papers associated with the ISMS-MSI Practice Prize). This work is related to a larger literature in marketing that has been investigating the value of conditioning promotion targeting on behavioral history (for representative papers, see, for instance, Rossi et al. 1996, Ansari and Mela 2003; for a review, see Arora et al. 2008). The most closely related within these are a subset of studies that use partial or full knowledge of the rules by which marketing is allocated across units to accommodate the reverse causality associated with nonrandom targeting. Here, this paper's empirical strategy can be thought of as an extension to a likelihood-based setting of a regression discontinuity–based strategy for identifying response under targeting (e.g., Busse et al. 2010, Hartmann et al. 2011). The advantage of this approach is that it utilizes the variation across all consumers within a bin for inference and is 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. The method can also be thought of as extending the Manchanda et al. (2004) approach for handling targeting to 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. 2017, Sudhir et al. 2016).

In 1970, John Little (1970, p. 1841) noted with concern:

> 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.

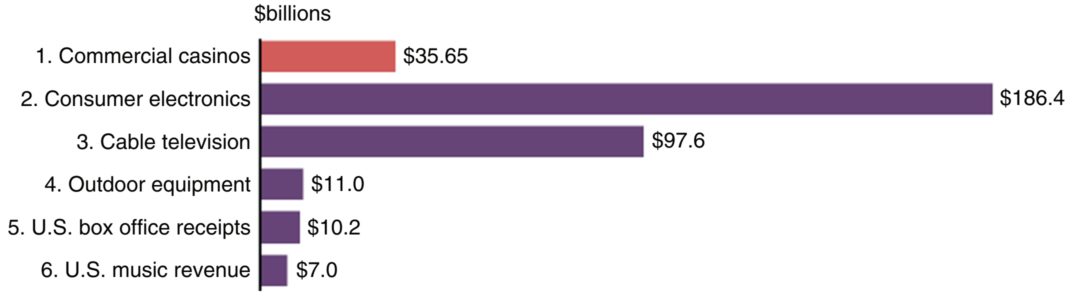
See Bucklin and Gupta (1999), Leeflang and Wittink (2000), Roberts (2000), Winer (2000), Lodish (2001), Sinha and Zoltners (2001), and Van Bruggen and Wierenga (2001) for various perspectives on the research practice divide. The availability of large quantities of consumer-level data and the increased recognition of the value of analytics provides for guarded optimism that the trend has turned in the other direction in this decade. One of the goals of this paper is to demonstrate in some detail the improvement from the adoption of an analytics-driven approach to marketing, in the hope that the productive collaboration between academia and industry will be accelerated.

The rest of this 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

Figure 1. (Color online) Comparing U.S. Gaming Revenues to Other Entertainment Spending (American Gaming Association 2012)



Sources. 1, American Gaming Association; 2, Consumer Electronics Association; 3, National Cable and Telecommunications Association; 4, Outdoor Industry Association; 5, Boxofficemojo.com; 6, Recording Industry Association of America.

place the size of the market at about $35.6 billion 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 (American Gaming Association 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 spending on gaming grew more quickly than that on any other component of the recreation sector (see Figure B.1 in Appendix B, reproduced from Bazelon et al. 2012). Clearly, gaming is an important part of the entertainment economy.

Within the commercial casino market, the state of Nevada accounts for about 30% of total revenues ($10.7 billion 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 off in Las Vegas in the 1940s with the opening of El Rancho 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 professional (Encyclopedia.com 2012). In 2011, the gaming revenues of the roughly 40+ casinos on the four-mile stretch of Las Vegas Boulevard known as "the Strip" alone accounted for approximately US$6.1 billion. This makes this area the top commercial casino market in the country. Figure B.2 in Appendix B depicts the casinos on the Strip. There is considerable agglomeration. The consolidation helps in demand aggregation, but also results in an intensely competitive environment for casinos on the Strip. More recently, casinos on the Strip have faced competition from the growth of international markets like Macau, from the gradual relaxation of gambling rules across states within the United States, and increasingly from online gambling outlets.

We now discuss some 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 United States are differentiated in scale and scope into two types, namely, destination casino resorts and regional casinos (Bazelon et al. 2012). Destination casino resorts are large facilities offering gaming, entertainment, retail, and convention facilities, and involve development costs that often exceed US$1 billion. Destination casino resorts attract visitors from all 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 are destination casinos.

Destination casinos are multiproduct 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 attract 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 subsidies 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 United States earned about 69% of their revenues from gaming, 13.2% from F&B, 10.4% from hotels 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 in locations like the Strip implies competition for consumers is intense. Hence, marketing becomes 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"), and discounts on food/drinks, as well as concierge services, 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 the marketing effort is targeted. As a general rule, more attractive promotions are offered to those who are expected to play more.

Targeting in the gaming context is a complicated problem. The extent of consumer heterogeneity is large, 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 sociodemographic 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 Las Vegas casinos have grown significantly in recent years, and many industry observers feel that much of comping does not drive incremental demand, being delivered to many without a measurable incremental effect on spending. While in the past comping was seen as a reward that had the effect of increasing play spending, now many believe some consumers see it as a precondition 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 the casino money. These issues have parallels in 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 for reasons analogous to those 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 is based on recency–frequency–monetary value (RFM) models, that preclude finer segmentation. Casinos also offer promotions in packages, bundling varying levels 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, and each 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 coordinated 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. Furthermore, there is a need to understand the impact of promotions at the property level because of 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 benefits to the firm.

Current Practice. Some aspects of current practice in targeting have been alluded to above. Many casinos are not analytically sophisticated in their marketing targeting practice and employ even more data-agnostic targeting rules compared to an empirically driven history-based strategy. The 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 in 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 the coin-in × 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 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 $\$100 \times 0.20 + \$100 \times 0.15 = \$35$. The 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, the actual win over the day is equal to $\$21 + \$13.5 = \$34.5$. The average daily theoretical (ADT) is simply the average *theo* over all of 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 because of 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 the average *theo*. More promotions are allocated to those that are observed to have higher realized *theo* wins in the trips over the last year. Once consumers are scored on the average *theo* plus 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 the "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 data-rich environments. Because of large investments in data-warehousing technologies and the widespread 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 the 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 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 that valuable consumers may drop in and out of campaigns and are not consistently targeted. Second, it does not help us understand how promotions drive value, for instance, by helping us understand whether promotions work by increasing visitation, 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 to MGM in the *long run*. For instance, a consumer may have wagered little in his first visit for a trip-specific reason, but yet may 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.

The model-based metrics developed here have the advantage that they use data on observed behavior from 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 visit. 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 enable 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 the development of a scalable, data-driven, microtargeting policy. This is, in essence, the new approach implemented at MGM.[5]

In the remainder of this paper, we describe the features of this approach, and 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. Does i visit an MGM casino in month t? Denote this by the binary variable $y^{(1)}_{i0t}$.
2. Conditional on visiting, does i visit property $j \in (1, \ldots, J)$? Denote this by the binary variable $y^{(1)}_{ijt}$.
3. Conditional on visiting, how much does i spend? Denote this by the continuous variable $y^{(2)}_{it}$.

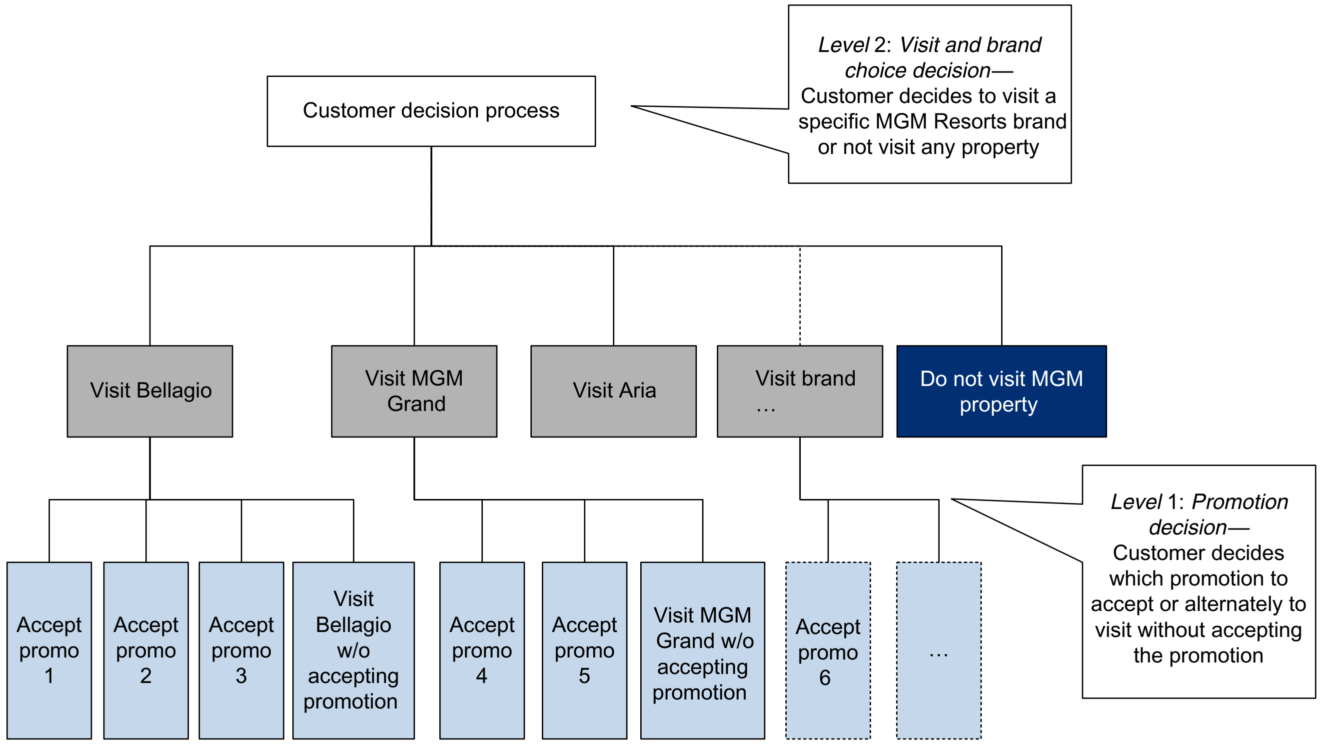
Collect these in a vector $\mathbf{y}_{it} = (y^{(1)}_{i0t}, y^{(1)}_{i1t}, \ldots, y^{(1)}_{iJt}, y^{(2)}_{it})$. Assume that there are K_{jt} different bundles of marketing promotions offered at property j in month t, and let x_{ikjt} be an indicator for whether the kth promotion bundle was offered 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., two tickets to a show + \$100 free-play valid only at the Bellagio, or a suite upgrade at any of the MGM properties). Collect the promotional offers valid for property j for month t in a vector $\vec{x}_{ijt} = (x_{i1jt}, \ldots, x_{ikjt}, \ldots, x_{iK_{jt}jt})$, and collect all the promotion vectors across properties for the individual in an array, $\mathbf{x}_{it} = (\vec{x}_{i1t}, \ldots, \vec{x}_{ijt}, \ldots, \vec{x}_{iJt})$. Let d_i be a vector containing the observed sociodemographics of consumer i. We fit a model of the form

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

where $f(\cdot)$ is a parametrically chosen link function (discussed below), and ϵ_{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 in that the model parameters, Ω_i, are allowed to be consumer specific. The goal of estimation is to learn the parameters Ω_i. The subset of Ω_i relating to the direct effect of $\mathbf{x}_{it}$ on $\mathbf{y}_{it}$ represents the effect of promotions on outcomes and is important to identifying a set of desirable consumers for subsequent targeting. The data for estimation include observations on $(\boldsymbol{y}_{it}, \boldsymbol{x}_{it}, \boldsymbol{d}_i)$ for a large sample of consumers (over 1 million) 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 $f(\cdot)$ for estimation.

Figure 2. (Color online) Nesting Structure Used in Model Setup



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, $y_{i0t}^{(1)}$, and the choice of which property to visit, $y_{ijt}^{(1)}$, as a standard 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 a visit to a property.[6] The self-selection of consumers into a promotion bundle is in and of itself informative of types (e.g., Chiang 1995), and we accommodate the information content of these choices into our estimation. We specify the lowermost nest of the discrete-choice model as a choice over the use of one (or none) of the offered promotion bundles. The higher-level nests then capture the choice of property or to not visit. Figure 2 depicts the nesting structure.

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

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

where, $\psi_{i\cdot}$ 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

$$\Pr(y_{ijt}^{(1)} = 1) = \frac{\exp(v_{ijt})}{1 + \sum_{j=1}^{J} \exp(v_{ijt})},$$

where the consumer-specific attractiveness of property j in month t, v_{ijt}, is specified as

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

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. Further, 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 j. $g(\cdot)$ 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 $g(\cdot)$ 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 $g(\cdot)$ 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

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

In the above specification, $h(\cdot)$ is a function of the current and past τ_2 trips made by the consumer, which allows for state dependence in spending. We allow $h(\cdot)$ to be a flexible linear function comprising the main and interaction effects of current and past visitation behavior, promotion utilization, and demographics, indexed by the parameter vector θ_i. Finally, 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 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 the models presented 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 for prediction alone. Segmentation requires understanding the 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 reflects an effect of those promotions on outcomes or the effect of outcomes on promotion allocation as induced by the targeting rule. In our setting, we find accommodating the reverse causality induced by targeting important to avoiding overstating the effect of promotions.

The approach below uses a partial (but not perfect) solution to the problem. To understand it, note that the 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). The concern is less about consistency of coefficients and more about producing a better out-of-sample prediction of promotional effectiveness that could ultimately be a reliable input to the formulation of an improved promotion policy. We will demonstrate later in this paper via a field implementation that the approach produces a better promotion strategy for the firm than the existing policy (or doing nothing).

To describe the setup, we use the notation in Equation (1), where the model is outlined in general terms relating outcomes $\mathbf{y}_{it}$ to $(\mathbf{x}_{i,t-\tau:t}, \mathbf{y}_{i,t-\tau:t-1}, d_i, \boldsymbol{\epsilon}_{it})$. This helps makes clear the complication induced by behavioral targeting. 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

$$\begin{aligned}
&\mathscr{L}(\{\Omega_i\}, \phi) \\
&= \prod_{i=1}^{N} \prod_{t=\tau+1}^{T_i} f_{y_t \mid y_{t-\tau:t-1}, x_{t-\tau:t}, d}(y_{it} \mid y_{i,t-\tau:t-1}, \mathbf{x}_{i,t-\tau:t}, d_i; \Omega_i) \\
&\quad \cdot \prod_{i=1}^{N} \prod_{t=\tau+1}^{T_i} f_x(\mathbf{x}_{it}; \phi),
\end{aligned} \tag{4}$$

where $f_{y_t \mid y_{t-\tau:t-1}, x_{t-\tau:t}, d}(\cdot)$ is the conditional density induced on $\boldsymbol{y}$ by $\boldsymbol{\epsilon}$, $f_x(x_{it}; \phi)$ 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

$$\begin{aligned}
&\mathscr{L}(\{\Omega_i\}, \phi) \\
&= \prod_{i=1}^{N} \prod_{t=\tau+1}^{T_i} [f_{y_t \mid y_{t-\tau:t-1}, x_{t-\tau:t}, d}(y_{it} \mid y_{i,t-\tau:t-1}, \mathbf{x}_{i,t-\tau:t}, d_i; \Omega_i) \\
&\qquad \cdot f_{x \mid d, \mathscr{H}}(\mathbf{x}_{it} \mid d_i, \mathscr{H}_{it}; \Omega_i, \phi)].
\end{aligned} \tag{5}$$

The likelihood no longer factors because $\mathbf{x}$ is set with some knowledge of the consumer's type Ω_i, his history $\mathscr{H}_{it}$, and characteristics d_i. Intuitively, 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 of Ω_i causing a first-order bias. Moreover, not knowing the true density of $f_{x \mid d, \mathscr{H}}(\cdot)$ 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 the analysis by Manchanda et al. (2004), but the difference is that their specification does not allow the density of $\mathbf{x}$ to condition on the consumer's history.

Our approach is facilitated by that fact that we know the variables on which targeting by the casino is based, namely, average *theo* and demographics, and 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

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

The likelihood has two parts, the first representing the conditional distribution of $x \mid z, \tilde{d}$, 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 $x \mid z, \tilde{d}$. 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 $\mathbf{y}$. Intuitively, as $\mathbf{x}$ changes, it produces both a direct effect on $\mathbf{y}$, due to the impact of promotions on outcomes, and an indirect effect by changing the set of individuals targeted. Only the first type of variation is useful for measuring the 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 are associated with changes in $\mathbf{x}$ while holding z and $\tilde{d}$ fixed are useful in learning about the direct effect of $\mathbf{x}$ on $\mathbf{y}$. Conditioning in this manner helps the model utilize that variation appropriately.

If we knew the density $f_{x\mid z,\tilde{d}}(\cdot)$ 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 $f_{x\mid z,\tilde{d}}(\cdot)$ 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 the 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 $f_{x\mid z,\tilde{d}}(\cdot)$. Therefore, we need a way to estimate the parameters 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 $z \mid \mathscr{H}$. 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, *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 by $\boldsymbol{y}^{(2)}_{i,t-N:t}$; the vector of expenditure *splits* of that outlay across the various available entertainment options on those past trips, denoted by $\boldsymbol{w}_{i,t-N:t}$; and 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

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

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 then is to model how the score is generated, by explicitly modeling the right-hand side 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 (because of 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 estimate the parameters without having to know the density of z.

Situations analogous to this problem are also faced by researchers. Often, 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 unmodeled, 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). For this reason, we cannot ignore the contribution of $f_{z|\mathscr{H}}(\cdot)$ to the full likelihood for estimating Ω_i.[7]

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

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

We can summarize the econometric difficulty as twofold. (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 a lack of knowledge of the process or because of incomplete information. Furthermore, the missing information is not random.

Exploiting Knowledge of the Targeting Rule. The solution to the issue here 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 heterogeneous and promotions are multidimensional, it is not necessarily the case that consumers in the highest *theo* segments are allocated to 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 to only a subset within the segment, rather than to all within it, is that the firm faces constraints that prevent it from blanketing everyone in the segment. For instance, the committee may face a limit on the number of promotions it can offer for stays at a particular property or for visits at a particular show because of 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.[8] 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 estimation.

First, we divide z into discrete bins and form segments by combining the bins of z and the bins of demographics used by MGM for targeting. Formally, letting $i_z\in(1,\dots,\mathscr{I}_z)$ denote the bins on the z dimension and $i_{\tilde{d}}\in(1,\dots,\mathscr{I}_{\tilde{d}})$ denote the bins on the demographics dimension, we define $R=\mathscr{I}_z\times\mathscr{I}_{\tilde{d}}$ 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 $r\in(1,\dots,R)$, 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 on only bin r, in which z takes values between $(\underline{z}_r,\bar{z}_r)$. Consider the subset of N_r consumers who have been assigned a priori to segment r, and let $\Omega^{(r)}$ represent the segment-r-specific parameter vector. Consider the likelihood for only this segment

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

The variation of $\mathbf{x}$ *within* $z\in(\underline{z}_r,\bar{z}_r)$ 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 a situation analogous 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, $f^{(r)}_{x\mid\tilde{d}}(\cdot)$, that does not depend on z, and therefore does not depend on $\Omega^{(r)}$. Hence, when estimating this model within segment r, we can ignore the part of the likelihood corresponding to $f^{(r)}_{x\mid\tilde{d}}(\cdot)$. Since we do not need to know $f^{(r)}_{x\mid\tilde{d}}(\cdot)$ for inference, this solves the first problem outlined above.

Second, note that 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.[9] This means the marginal density of z, which determines segment membership by specifying the probability that $z \in (\underline{z}_r, \bar{z}_r)$, 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, we divide observations into R nonoverlapping segments in a first step, and then estimate separate models for each subsegment. The idea 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. More generally, behavioral targeting by firms results in a complicated selection on unobservables problem in estimation. Observing the variables on which targeting is based on converts the problem of selection on unobservables to one of selection on observables, which facilities control for nonrandom selection. In practice, we expect this method to be attractive because it is simple to implement and exploits the internal information of firms, which is usually available in such settings. A useful feature is that, since the likelihood of the score can be ignored in estimation, researchers need not know exactly the factors that determined the historical evolution of these metrics in the firm's database to use this method. The policy of assigning marketing on the basis of bins of summaries of historical behavior (like RFM value or other metrics) is common in industry. Hence, we expect this to have some applicability in other contexts.

A caveat to this analysis in our setting is that 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 that 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, the implementation here should be seen as approximate. The method will be exact in other situations where researchers know the bin cutoffs more precisely. Some sensitivity to this is assessed in the simulations reported in Appendix A, which also briefly discusses factors that determine the applicability of this procedure in other settings.

Parameter Interpretation and Prediction. A feature of the binning method above is that it produces z-specific as opposed to individual-specific parameters. What it is doing is characterizing an individual's type by his z and $\tilde{d}$, and then obtaining type-specific parameters. Thus, z and $\tilde{d}$ become the relevant dimensions of heterogeneity. One way to think of this is as a model of time-varying heterogeneity, in which the variation over time in an individual's parameters is projected onto changes in a summary measure of his historical behavior, z. A restriction is that the variable onto which heterogeneity is projected has to be the variable the firm uses for its segmentation. Therefore, the richer the range of the firm's segmentation policy, the richer the range of heterogeneity the researcher is able to accommodate in this way. 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.

Once the segment-specific estimates are obtained as above, we predict the response in the following way. For individual i, we find his current average *theo* (z_{it}) and demographics ($\tilde{d}_i$). The values of these two sets of variables together determine which bucket r consumer i 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 the lowest cost.

5. 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 each consumer in each month in the data), (2) matching the different sources of information based on unique identifiers (e.g., matching consumer IDs in the transaction database to consumer IDs in the marketing databases of corporate and property-specific departments), and (3) cleaning the data to eliminate database coding errors, unreasonable entries, and/or missing information.

Data Descriptives. The data consist of individual-level transactions of a random sample of about 1 million consumers 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 with marketing offers, some are not. Visitation and transactions of all consumers are observed, as are details of all of the offers mailed out and redeemed. Most 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. 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 that the analytics solution needs to take as given. Developing a database that tracks consumer behavior across competing casinos will be an important step forward for the industry as a whole to better capture competition and "share-of-wallet."

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 because of business confidentially concerns.) These segments are 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 plays at slots or tables.[10] Table B.1 in Appendix B shows the R segment definitions, as well as the proportion of consumers and the number of trips observed in each bin.

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

Past history $\{g(\cdot)$ *and* $h(\cdot)\}$. To operationalize the functions $g(\cdot)$ and $h(\cdot)$ capturing the history of past play, we include several metrics of past history, including average bet, coin-in to point 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 the *theo* and actual win at tables and slots separately, in residential casinos versus nonresidential casinos, and in luxury versus nonluxury properties (based on MGM's definitions of property classifications); the *theo* and actual win arising from free play-based play; and other metrics of dollars of incented versus nonincented play.

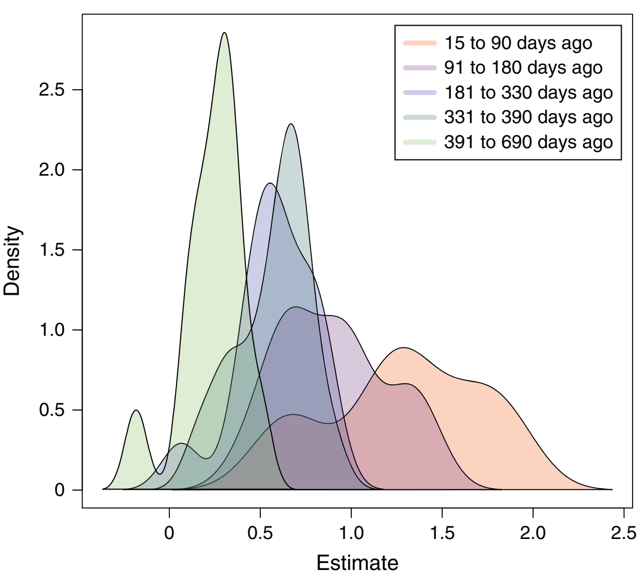
Marketing offers. We include variables measuring the range of marketing offers from the company at both the corporate (valid at all properties) and property levels. These offers comprise a large set and include the following:

1. *Room* metrics like room type, room discount, number of comp nights, and whether the comp is midweek or weekend
2. *Entertainment, sports, and facility* offer metrics like club pool offers, entertainment type, an indicator for entertainment offers, ticket price discount, an indicator for facility offers, an indicator for sports offers, sports offer amounts, sports ticket price discounts, and an indicator for golf offers
3. *Casino event information* metrics like an indicator for inclusion in the casino event prize pool, the prize pool format, an indicator for grand prize inclusion, the grand prize format, the prize value offered, cost of the event for which the offer is made, buy-in amount, points to entry if offered, and tier credits to entry if offered
4. *Special event* metrics like indicators for special events, 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 an indicator for a retail offer, retail offer amount, an indicator for a spa offer, and spa service amount
6. *Air and limo* offer metrics like an indicator for an airline offer, the air package amount, an indicator for a limo offer, and an indicator for a 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

6. Results

Estimation is programmed in the SAS statistical package. The final implementation involves about 120 separate estimated models of consumer behavior (separated by segment, casino, and outcomes), about 180+ variables in each model, and about 20,000 parameters estimated across these models. We present a representative set of results for brevity. We report parameter estimates and not marginal effects for business confidentiality reasons. Figure 3 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 3 presents these distributions. Looking at Figure 3, 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 ago (pink distribution) are, on average, roughly six times more likely to visit than those who have visited more than 391–690 days ago (lime green distribution). This may also reflect within-segment heterogeneity in that the first bucket comprises consumers with high utility from gambling and visitation, while the second reflects those with lower value (or costs) from visits. These duration effects allow the model to link a customer's visitation behavior over time in assessing his relative value to the casino. For 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 3. (Color online) Effect of Time Since Last Trip on Visit Propensity

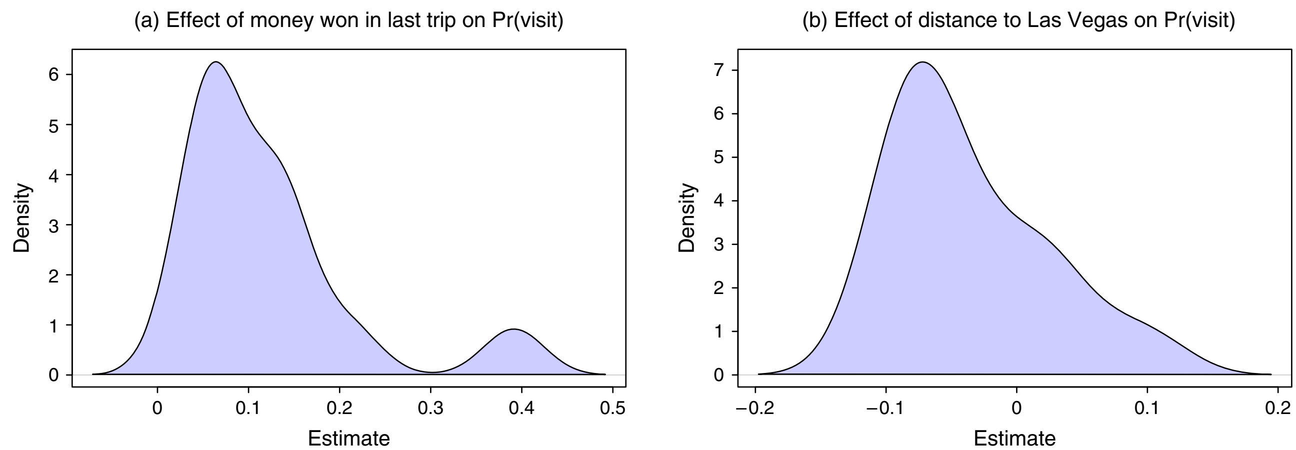


Note. Histograms with more mass on smaller values correspond to larger *time since last trip*.

Figure 4(a) 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 4(a) plots the density of the coefficient on this variable across the R segments. There is an interesting heterogeneity in the effect of past winnings on current visitation: there is a bimodal 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 of free-play promotions, for instance. Figure 4(b) 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 a variation in distance across consumers within each segment). Figure 4(b) plots the density of the coefficient on this variable across the R segments. Interestingly, we find that 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.

As a representative example of the effect of targeted marketing, we plot the effect of providing a free room on visitation. We estimate a separate effect of a free 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 5 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 5, 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, to result in better optimization of promotions across customers and properties in a subsequent stage.

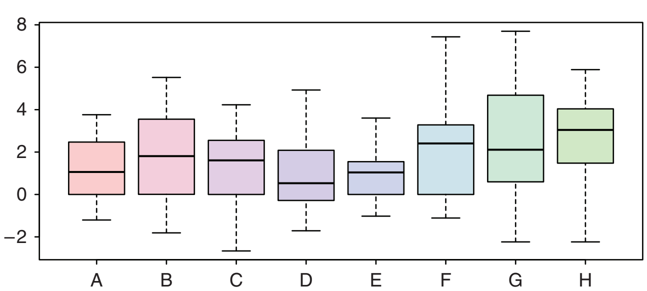
Figure 4. (Color online) Effects on Visit Propensity



Finally, we present plots of the effect of customer characteristics on spending conditional on visit. Figures 6(a) and 6(b) 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 Figure 6(a) 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 6(a) shows the distribution across the R segments of the effect of being a male aged 25–35 relative to a male aged < 25 years on spending propensity. Figure 6(b) shows the analogous plot for females. Interestingly, we see few systematic differences in spending, all things held equal, across various age tiers for males. However, the distribution is an inverted-U-shape for females: women aged 25–35 are significantly less likely to spend compared to those below 25 years old, 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.

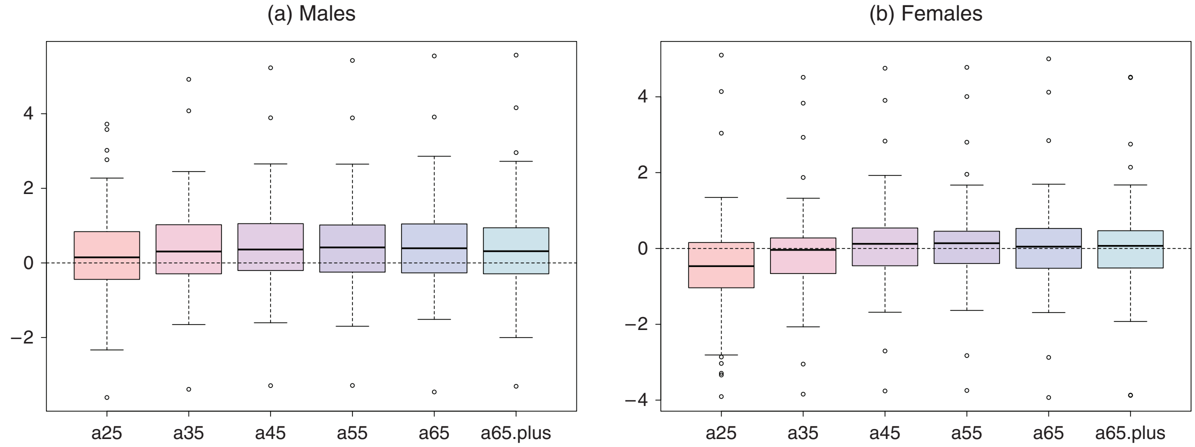
Figure 5. (Color online) Effect of Free Room (Relative to Property X)



7. Optimization

The second module involves an optimization platform that searches within a specified promotion set to recommend a promotion bundle for each consumer in the database. The optimization is implemented at the campaign level and is operationalized in the following manner. First, the corporate marketing team managing the campaign decides a set of component promotional options that could potentially be offered to the customer base (e.g., a given level of discount at a new property or a given level of free-play credits). This decision is outside of the optimization and reflects the goals the management of the company has in running the advertising campaign (e.g., drive visitation at a new property, increase play at slots, etc.). Taking these component promotional options as given, the optimization package scores each customer on each possible component or bundle of component promotions. The scoring is based on the models of visitation and spending outlined above. To do this, it first calculates, for each customer, the expected spending (unconditional on a visit) minus the expected cost to the firm of offering each promotion component/bundle combination to the customer. Then, it calculates, for the same customer, the expected spending (unconditional on a

Figure 6. (Color online) Effect of Customer Age on Spending



visit) in the absence of offering those promotion components/bundles. The difference between these represents the incremental expected benefit to the firm of offering that customer a specific promotion component/bundle combination. If the expected benefit is negative for all promotion component/bundle combinations, that customer is sent no promotions. If the expected benefit is positive, the promotion component/bundle combination that yields the highest value is allocated. We also incorporated into the optimization platform the flexibility to add more decision rules if desired by management. For example, management could 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 for a customer to be eligible.

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 Section 8 involves scoring about 1.5 million consumers on 75+ promotional options. The optimization package should also integrate well with the statistical models presented above and provide managers the ability to add constraints to the optimization in a user-friendly manner. We implemented these on Teradata, a commercial database platform 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 at the individual consumer level, facilitating finer microtargeting. Furthermore, compared with the prior system, the number of bundles that could be considered increased by about six times, increasing the number of instruments available to improve promotion efficiency.

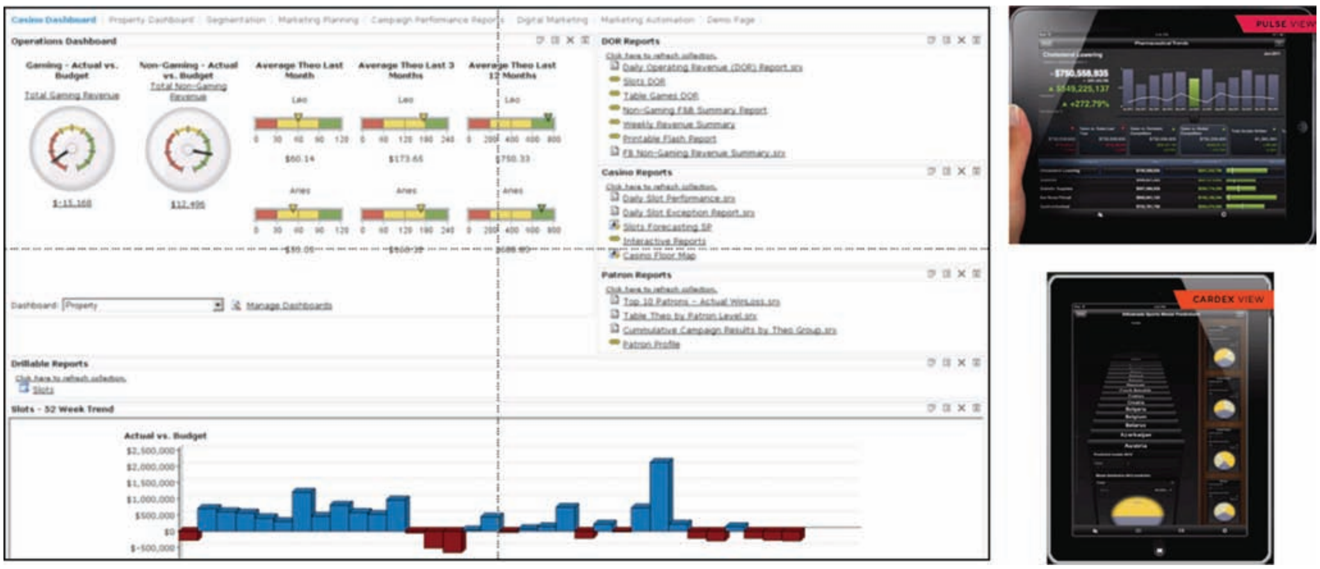
Dashboard applications were also developed that enabled managers to monitor and dissect company performance on their desktops. Figure 7 provides an example. These dashboards were linked to the underlying statistical model and optimization packages, to embed the framework in a user-friendly decision support system. This completes the discussion of the model framework and development effort.

8. Results from a Randomized Evaluation

The model developed above was assessed as part of a randomized controlled test at MGM. The test evaluated 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 relative to the status quo approach used at the firm.

Setup. First, a pilot test was conducted in the spring of 2012 with a limited number of promotional offerings to assess the test design, understand ballpark customer response, and understand the logistics of implementation. Based on this, the test was implemented as part of the summer 2012 corporate campaign. To do this, managers from MGM's corporate marketing team first created 75+ promotion bundles that comprised the superset of possible corporate promotions that could be offered to consumers in the campaign. Two million consumers from MGM's prospect pool for such campaigns were picked and randomly divided into three groups. Customers in Groups A and B were then allocated promotions taking the superset as given. Group A consumers (30% of the total) were allocated promotions according to the model's recommendations, and Group B consumers (30% of the total) were allocated promotions based on the status quo approach (past *theo* and demographics). Group C consumers (10% of the total) were treated as the control—they did not receive any corporate offers. The remaining 30% of consumers

Figure 7. (Color online) Screenshots of Dashboard Applications for Campaign Management and Monitoring Linked to Underlying Empirical Models



were tested on auxiliary aspects that were unrelated to the model and not relevant to the evaluation. All allocated corporate offers were then emailed to consumers before the beginning of the summer. The offers were valid for redemption from July 31, 2012 to October 31, 2012. All visits to any of the MGM properties along with all transactions involving any of the 1.5 million consumers in the test were then tracked during the July 31–October 31 window during which the promotion was active. No other corporate marketing was targeted at these consumers during that time.

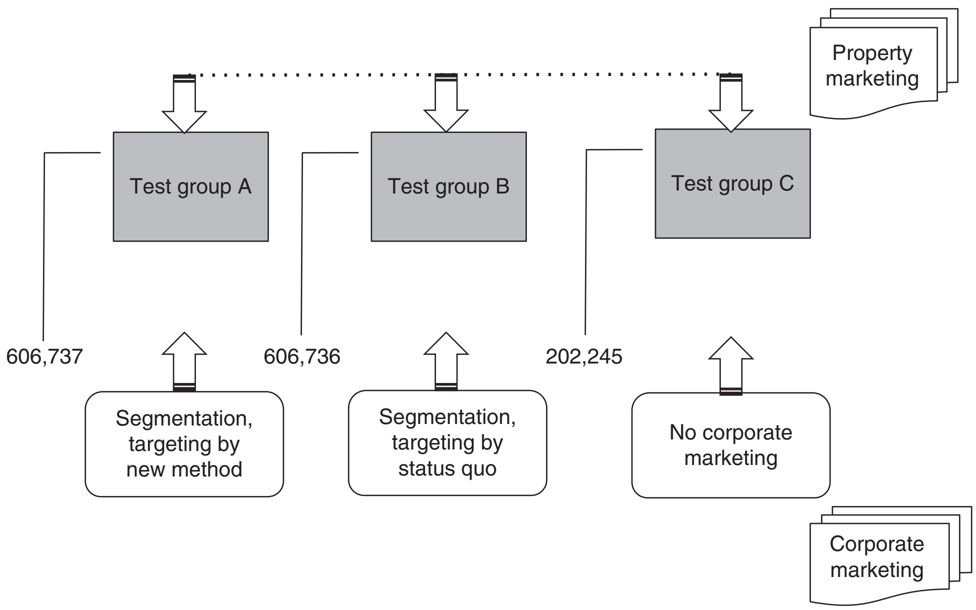
Allocation Process. To allocate the promotions, all customers in Group A were scored by the optimization package on each bundle in the superset. For each bundle, the incremental expected benefit to the firm of offering the customer that bundle was calculated as described in the previous section. A restriction was imposed that each consumer could be offered only one bundle during the campaign. Then, each customer was assigned the promotion bundle that provided the highest incremental expected benefit subject to it clearing a minimum threshold. The threshold was set outside the model by management. If the bundle with the highest incremental expected benefit did not meet the minimum threshold, no corporate promotion was sent to that consumer. In group B, consumers were sorted into segments as per the status quo method. Managers then assigned promotion bundles to each segment in the group in the same way as they have made those decisions in the past.

Concurrent Property-Specific Promotions. During the same period, the three groups were also exposed to *noncorporate* campaigns set independently by the MGM resorts' individual property marketing teams. Thus, those in the control and treatment groups received other property-specific promotion offers over and above those associated with the corporate test. This aspect 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. The promotions for the individual properties were mailed out in the summer prior to July and were set independent of the corporate campaign, so they were not jointly allocated or adjusted in response to the corporate-level interventions in Groups A and B.[11] Figure 8 presents the test design pictorially.

Performance Metrics. To assess the performance of the test, we report revenues and cost metrics for the three groups, as well as an ROI measure. We report these, as these are the metrics that are used for decision making at the firm. For business confidentiality reasons, we scale all dollar numbers by an undisclosed constant, so all numbers we report 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.

As our revenue metric, we report an "adjusted" value for each group. These values are constructed by summing all gaming and nongaming *theos* from agents in a group that visited during the July 31–October 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. This makes the adjusted revenues a more conservative metric of the gains from a campaign.

Figure 8. 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, and Group B consumers were offered corporate promotions based on the status quo method. Group C (Control) consumers were offered no corporate promotions. All groups continued to receive promotions offered by the individual properties. The offers were valid for redemption from July 31, 2012 to October 31, 2012, and were mailed out in the summer of 2012. All visits to any of the MGM properties along with all transactions involving any of the 1.5 million consumers in the test were then tracked during the July 31–October 31 window, during which the promotion was active.

Computing the right cost metric is more difficult. The economically relevant cost for MGM is the opportunity cost of allocating a promotion bundle to a customer. This requires computing the next best option for a promoted product had it not been offered to a given consumer, and what the consumer would do in the absence of receiving the promotion. These are difficult constructs to evaluate in this complex setting. Given this, we report the dollar value of the promotion as the measure of cost. This is the metric used at the firm as well. While this metric maps into how the firm evaluates these results, one caveat is that it does not reflect exactly the true economic cost of the promotion. This is a limitation of this analysis, one that is shared with the existing broader literature on promotions.[12] Reported costs for Groups A and B refer to the costs incurred by MGM via redemption of either corporate promotions or property-specific promotions assigned to consumers in that group. Because of reporting constraints, we are unable to split these out separately by corporate versus property-specific redemptions. The costs for the control group refer to the costs incurred by the properties to run other property-specific campaigns they conduct in parallel with the focal corporate promotion. Other costs of running the campaign (e.g., time costs of sending the email) are small and are not included.

Finally, to assess performance, we report an incremental ROI metric defined as adjusted revenues over costs of the new policy relative to the status quo and the control; as well as incremental profits defined as adjusted revenues minus costs relative to the status quo and the control.

Interpretation of the Test. The *treatment effect* of the test should be interpreted as follows. The inputs to Groups A and B are a superset of ex ante identical promotions and consumers. The treatment is an allocation from the set of promotions to the set of customers, taking these as given. The posttest difference in metrics between the two groups reflects a treatment effect. The difference between Group A and Group B represents the treatment effect of the model/academic–based allocation relative to the status quo–based allocation for a given set of promotions. The difference between Group A and the control group represents the treatment effect of the model/academic–based allocation for a given set of promotions, relative to the existing property-specific marketing activity. The difference between Group B and the control represents the treatment effect of the status quo–based allocation for a given set of promotions, relative to the existing property-specific marketing activity.[13]

The treatment effect of the model/academic–based allocation relative to the status quo is likely to be a lower bound for two reasons. First, further gains could be realized if the model were used to optimize the superset of promotional options in addition to its allocation to a pool of customers. Second, further gains beyond those assessed in the test could be had if property-specific campaigns were coordinated with the corporate campaign. Additionally, to the extent that the new method is better, implementing it at the individual-property level could be the source of additional gains from the adoption of the new analytics solution.

Results. Table 1 reports the results. In Table 1, the "Adjusted revenues" row for each group reports the total money made from each ex ante identical group from application of the column-specific treatment. The "Costs" row for each group reports the total costs incurred by each ex ante identical group from application of the group-specific "treatment." Looking at Table 1, we see that adjusted revenues from those treated under the status quo policy amounted to about R$111.97 million, compared to R$114.06 million under the new model. Thus, adjusted revenues are higher under the new policy. We also see net costs are about R$41.42 million under the new policy versus R$43.90 million 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$72.64 million profit to the casino under the new policy compared to about R$68.07 million under the status quo. The difference is about R$4.57 million 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 million and $5 million 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 1, we see that various other campaigns run concurrently by the individual properties brought in about R$36.8 million of revenue from consumers in the control group. Recall that the control group is one-third the size of Groups A and B, so to obtain a relative comparison, we should multiply the dollars in the control by 3. Computing scaled revenues 3 × R$36.8 million = R$110.4 million, 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$3.7 million more (R$114.1 million for the new policy versus R$110.4 million for the control). Computing costs, the individual properties spent a scaled total of 3 × R$14.5 million = R$43.5 million. The net profit impact is 3 × R$22.2 million = R$66.6 million, which is less than the R$72.6 million 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 any one property's method.[14]

To see these numbers on a per-consumer basis, we compute the relative revenues, costs, and profits for Groups A and B, scale these by the number of consumers, and report these on a per-consumer basis in Table 1. As an example, to compute the relative revenue on a per-consumer basis for Group A (reported in row "ΔRevenue per consumer (relative to Group C)," column "Group B" of Table 1), we take the adjusted revenue for Group A, R$114.06 million, subtract out the adjusted revenue for the control, 3 × R$36.77 = R$110.31 million, and divide by the number of consumers in Group A = 606,736, to get R$6.2 million.

Table 1. Aggregate Performance of Treatment and Control Groups

	Group A: New ($N = 606{,}736$)	Group B: Status quo ($N = 606{,}737$)	Group C: Control ($N = 202{,}235$)
Adjusted revenues (million)	R$114.06	R$111.97	R$36.77
Costs (million)	R$41.42	R$43.90	R$14.53
Margin (%)	63.68	60.79	60.49
Profit (million)	R$72.64	R$68.07	R$22.24
ΔRevenue per consumer (relative to Group C)	R$6.20	R$2.75	—
ΔCosts per consumer (relative to Group C)	−R$3.55	R$0.53	—
ΔProfit per consumer (relative to Group C)	R$9.75	R$2.22	—
Revenues ÷ costs	2.75	2.55	2.53
Incremental return on investment of new policy			
Relative to status quo		0.20	
Relative to no corporate marketing		0.22	

Notes. Group A consumers were offered corporate promotions based on the model, and Group B consumers were offered corporate promotions based on the status quo method. Control consumers were offered no corporate promotions. All groups continued to receive promotions offered by individual properties. The cost of the campaign is calculated as the net value of promotions offered. Other costs of running the campaign are not included. The "costs" row for Groups A and B 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 refers to the costs incurred by the properties to run the other campaigns they conducted in parallel to the focal corporate promotion. The ΔRevenue, ΔCosts, and ΔProfit numbers in Groups A and B are calculated by taking the adjusted revenues, costs, and profit numbers for each group, subtracting out the corresponding values for the control, and dividing by the number of consumers in each group. Incremental ROI is calculated as adjusted revenues/costs of the new policy relative to the status quo or to the control (no corporate marketing).

Doing this for Group B relative to Group C provides the relative revenue on a per-consumer basis for Group B ((R\$111.97 − 3 × R\$36.77) ÷ 606,737 = R\$2.75). Doing the same for costs and profits produces the two rows below.

Looking at these numbers, we see more clearly the source of the improvement of the new policy—relative to the control, the new policy makes more revenue at much lower costs. The status quo policy also brings in more revenue than the control, but at a higher cost. Comparing the new policy to the status quo, the revenue improvement of the new policy relative to the control is of the order of 2.3 times that of the status quo (R\$6.2/R\$2.75). These results also allow us to informally assess the underlying drivers of the improvement. Improvements from allocation can arise from two factors, one from the better matching of promotion types to household preferences implied by the model, and the other from the reallocation of promotions from average to marginal consumers (who are more likely to respond to the promotion). If the new model only reallocated the same promotions as before to consumers who were more likely to respond to them, we would have seen redemption costs go up (or weakly remain the same) and revenues weakly increase. The fact that we see costs go down and revenues go up suggests that better matching plays an important role in the profit improvement in addition to reallocation. Finally, in terms of profits, the new policy makes an incremental profit of R\$7.53 (R\$9.75 − R\$2.22) on a per-consumer basis (or between \$2 and \$4 in real dollars) relative to the status quo.

Finally, computing an incremental ROI, we find the new policy provides a net incremental ROI of about 0.20 compared to the status quo policy.[15] Thus, a dollar spent in promotions generates about 20¢ of incremental profit from consumers under the new policy compared to the current practice at the firm. As another metric, if the four campaigns in one year each spent the same amount on promotions as the summer campaign, at these levels of ROI, the firm would make about R\$33.6 million (or between \$10 million and \$15 million in real dollars) in incremental profit from using the new model compared to the status quo method, or about R\$41.7 million (or between \$14 million and \$19 million in real dollars) in incremental revenues compared to the aggregation of the campaign planning strategies of the individual properties.

9. Conclusions

Efforts on developing and implementing a marketing analytics solution for a real-world company are presented. The framework leverages the company's customer-level data to develop empirically informed 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 choice, and the controls 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 assessed relative to the status quo using a field intervention that shows that the profits from adopting the new system are high.

For academics wishing to port marketing science models from theory to practice, we discuss a few takeaways from this case study. First, one takeaway is well known in the literature, but is worth reiterating—capturing the heterogeneity within the response model is key to the targeting problem, and to do so effectively, a large number of segments is required. Even though we allow for $R = 50+$ segments, we could detect a significant amount of within-segment heterogeneity, some of which we do not model 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). What we have also tried to highlight in this paper is the usefulness of an approach to dealing with the heterogeneity, in which the semiparametric partitions used by the econometrician coincide with those used by the firm for targeting. In addition to handling heterogeneity, this helps address some concerns associated with reverse causality, even when those partitions are based on past behavior that has been influenced by historical targeting.

Second, we found that thinking structurally (i.e., What is the model that generates the observed variables in the data?) is useful in assessing the historical variation in the data and in using it to formulate policy. The structural approach requires the researcher to write down a fully specified model that generates the data. The reason this is useful even if the full model is not required to be estimated in a project is that thinking this way incentivizes the researcher to at least contemplate what generates the variation in *both* y and x. In the current setting, attempting to write down a model for how x (marketing) is generated both made clear the problem induced by historical targeting and helped suggest the solution to the problem by uncovering the within-segment targeting rules employed at the company. Accommodating the firm's historical targeting rule in inference was important to measuring the right effect of promotions, and for guarding against recommending a ramping up of promotions when using the estimated parameters for formulating marketing policy. Separately, attempting to write down a model for x and for the variables that drive x (like Equation (6)) showed the likelihood contributions of these components were not ignorable in the estimation of the parameters that explained the data, which led to the utilization of a method that did not require taking a stand on these terms. This simplified the analysis. Overall, we believe that while estimating a full set of equations describing "the entire system" along with associated "deep structural parameters" may not be necessary in many practical applications, specifying a full model may still be helpful in understanding how to approach the data and suggesting ways to interpret them correctly empirically.

A third aspect is that 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 an unstructured or nondigital 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 high importance and reaps large benefits because the value of the subsequent modeling is driven to a high degree by the quality of data inputs.

Finally, one limitation of the current paper is that several details of the data and estimates cannot be revealed because of business confidentiality agreements. Going forward, the question of how to structure arrangements such that the needs of confidentiality and academic transparency are balanced in a reasonable way will be an important one for the academic community to address.

Acknowledgments

Some numbers in this paper have been scaled or disguised to preserve confidentiality. The views discussed here represent those of the authors and not of Stanford University or the University of Chicago. The usual disclaimer applies. For helpful comments and suggestions, the authors thank participants at the 2014 Big Data Marketing Analytics Conference at the Chicago Booth School of Business; NASSCOM InnoTrek 2014; the 2014 Summer Institute in Competitive Strategy at the University of California Berkeley; the 2014 Successful Applications of Customer Analytics Conference at the Wharton School; the 2014 Theory and Practice Conference at the Kellogg School of Management; the 2014 Workshop on Social and Business Analytics at the University of Texas at Austin McCombs School of Business; the 2015 American Marketing Association Winter Conference in San Antonio, Texas; the 2015 Marketing Science Institute Regional Knowledge Networking Event in San Francisco; and the 2015 Google–Marketing EDGE I-MIX Program in Mountain View, California; seminar participants at Columbia University, the University of Delaware, the University of Iowa, Stanford University, and the Wharton School; discussants Vineet Kumar and Puneet Manchanda; and the journal editorial team, including Preyas Desai, J.-P. Dubé, Pedro Gardete, Chris Nosko, Thomas Otter, Stephan Seiler, Jeff Zweibel, and especially Gary Russell, Josh Swissman, and Eric Bradlow and Gunter Hitsch in particular.

Appendix A. Intuition for Estimator and Monte Carlo Studies

This appendix presents a linear example that illustrates the intuition for how the estimator works, and then reports on Monte Carlo simulations that investigate its performance. The simulations also assess 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 the empirical application. Finally, we briefly discuss some considerations that determine the use of this method in other settings.

Intuition. 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. To see why this is helpful, consider a cross-sectional model. Let i denote an individual, and suppose outcomes y_i, marketing x_i, and score z_i are generated in the following way:

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

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. To fix ideas, note 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 setup, x is endogenous because of 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 $z_i = 0$, $x_i = g_0(\nu_i)$, estimate $y_i = \alpha_{(0)} + \beta_{(0)} x_i + \epsilon_i$.
- When $z_i = 1$, $x_i = g_1(\nu_i)$, 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 because of 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 $g(\cdot)$ 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 this paper. While our approach has a 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.

Monte Carlo Study. The simulations below use the linear model above, though we modify the setup slightly for ease of illustration of some of the econometrics driving the results. We start by assuming the true data-generating process is as follows:

$$y_i = \alpha + \beta x_i + \epsilon_i. \quad \text{(A.1)}$$

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 $k = 1, \ldots, K$ that his score, z_i, falls into, as

$$x_i = \theta_0 + \sum_{k=1}^{K} \theta_k \mathbf{I}(z_i \in \mathcal{Z}_k) + v_i. \quad \text{(A.2)}$$

We model the process generating z_i as

$$z_i = \omega \eta_i + (1 - \omega) y_{i0}, \quad \text{(A.3)}$$

and assume that

$$\epsilon_i = \kappa(\eta_i; \varrho) + \varepsilon_i, \quad \text{(A.4)}$$

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 (A.3), ω is a weight such that $\omega \in (0, 1)$, and in Equation (A.4), $\kappa(\eta_i; \varrho)$ are nonlinear basis functions of η_i indexed by parameters ϱ. The parameters ω and ϱ 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 accentuate the endogeneity.)

We make the following distributional assumptions: $\eta_i \sim \mathbb{U}(0, 1)$, $y_{i0} \sim \mathbb{U}(0, 1)$, $v_i \sim \mathbb{N}(0, 1)$, $\varepsilon_i \sim \mathbb{N}(0, 1)$, 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 nonoverlapping splits of the unit segment. Our simulations vary the parameter vector $\Psi \equiv \{\alpha, \beta, \boldsymbol{\theta}, \varrho, \omega\}$, 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 specifications for the purposes of discussion. For most of our simulations, we also assume that $\kappa(\eta_i; \varrho) = \varrho_1 \eta_i + \varrho_2 \eta_i^2$.

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

$$y_i^{(k)} = \alpha^{(k)} + \beta^{(k)} x_i^{(k)} + \epsilon_i^{(k)}. \quad \text{(A.5)}$$

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

$$\Delta = \frac{\hat{\beta} - \beta_{\text{truth}}}{\beta_{\text{truth}}}. \quad \text{(A.6)}$$

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

$$\Delta^{(k)} = \frac{\hat{\beta}^{(k)} - \beta_{\text{truth}}}{\beta_{\text{truth}}}, \quad \text{(A.7)}$$

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

Results: Base Case. We first demonstrate that our approach recovers β with zero average bias across many replications. In each replication, a true Ψ vector is drawn randomly, and a corresponding data set of $N = 10{,}000$ units is generated by simulating from Equations (A.1)–(A.4) under the assumption that there are $K = 3$ bins. The simulated data set 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 (A.5) is estimated separately for each of the K bins in the simulated data set. Figure A.1 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 take only 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 reduce bias.

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

Figure A.1. Bias in Recovery of β Across Replications $\{N = 10{,}000, R = 100\}$

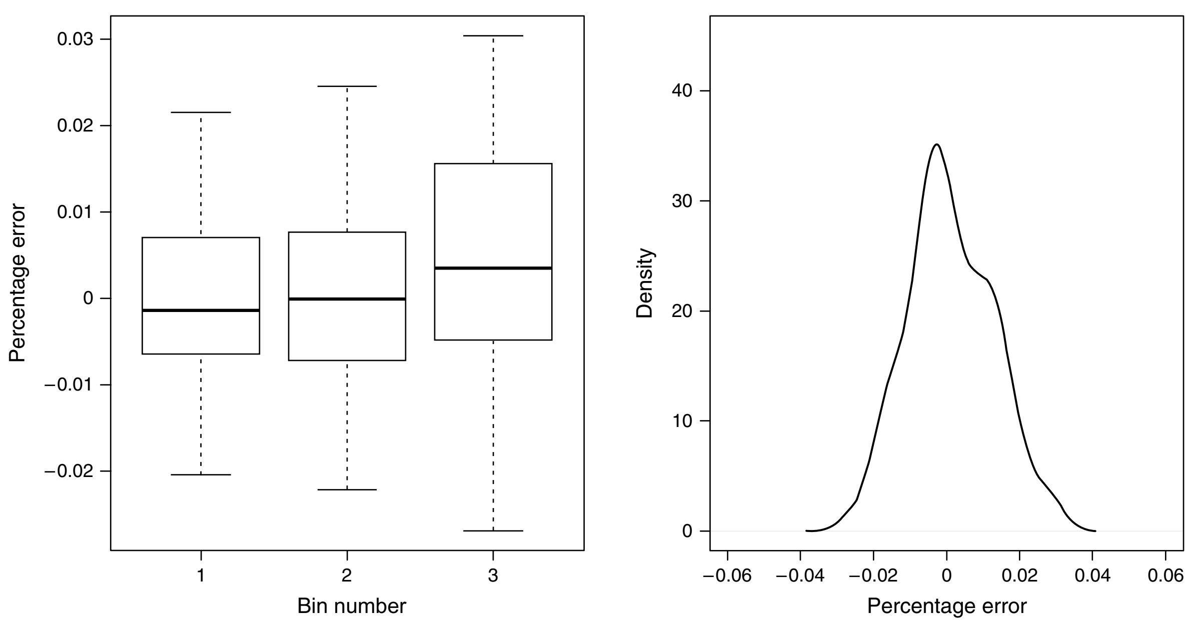
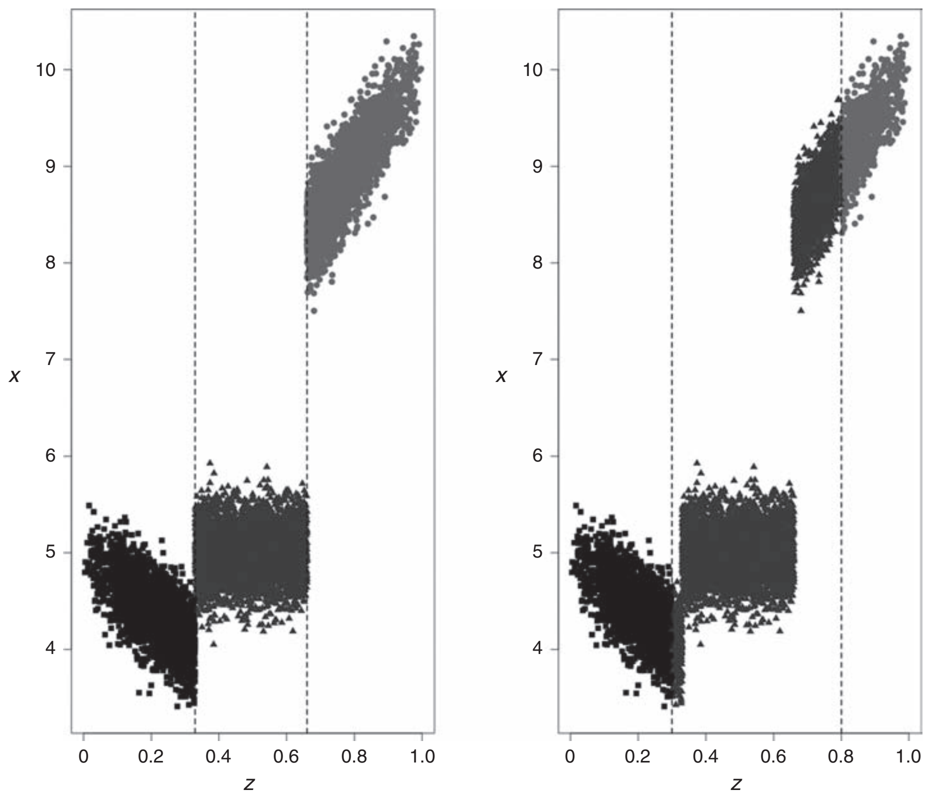


Figure A.2. Marketing (x) as a Function of Score (z) with Nonrandom Within-Bin Allocation and Misspecified Cutoffs



in the construction of the bins (i.e., $\mathcal{Z}_k$) and (b) when x_i is potentially nonrandomly 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

$$x_i = \theta_0 + \theta_1 z_i I(z_i \in \mathcal{Z}_1) + \theta_2 I(z_i \in \mathcal{Z}_2) + \theta_3 z_i I(z_i \in \mathcal{Z}_3) + \delta y_{i0} + v_i. \tag{A.8}$$

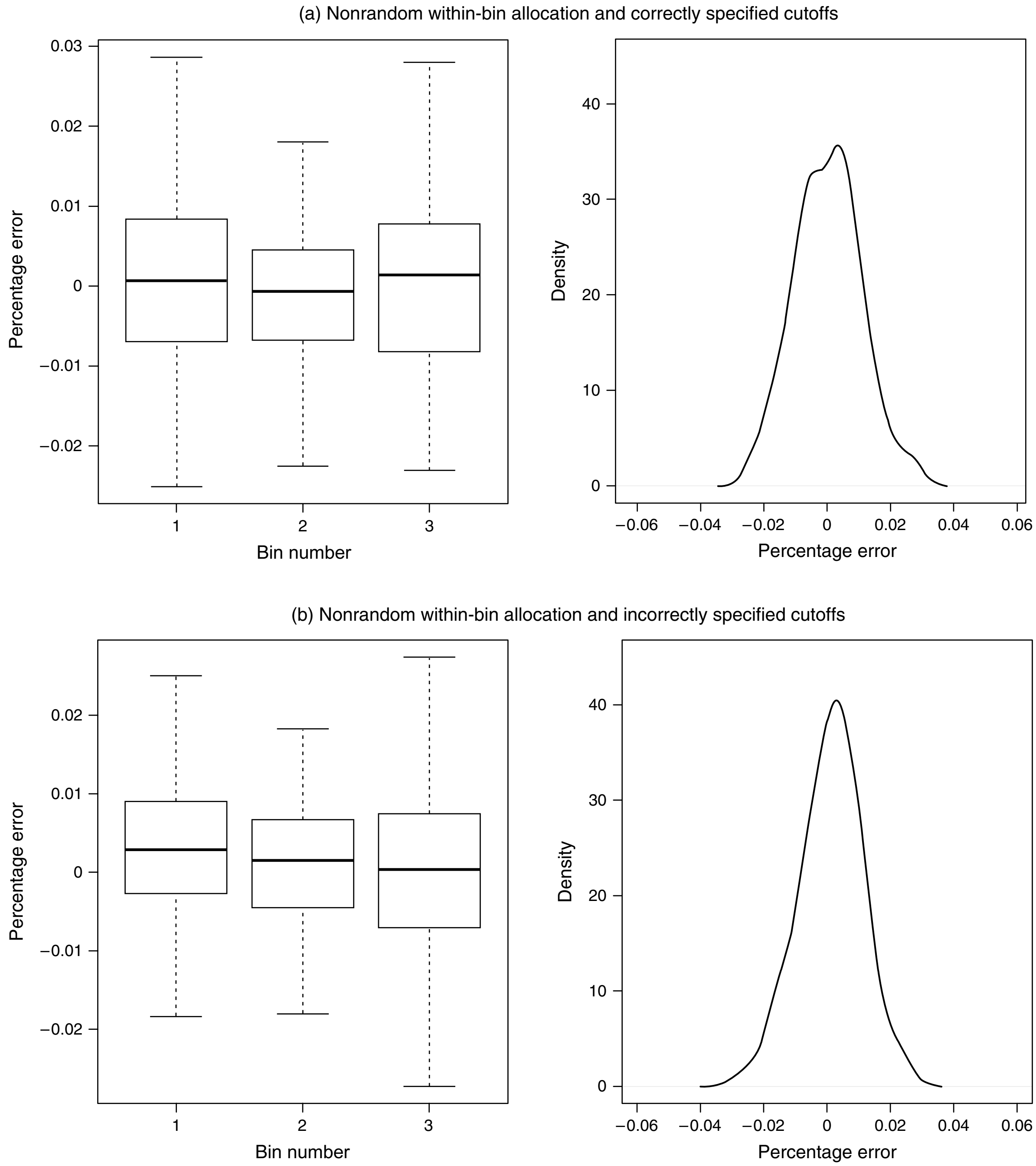
Note that in Equation (A.8), we do not allow for a direct dependence between x and z in the second bin. We do this to illustrate visually (below) the effects of cutoff misspecification; allowing the direct dependence in all bins does not alter the qualitative nature of the results below in any substantive fashion. We simulate $N = 10{,}000$ units from Equation (A.8) along with (A.1), (A.3), and (A.4) as before and plot the simulated data in Figure A.2. The left panel of Figure A.2 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 has misspecified the cutoffs. Looking at the left panel, we see that x is strongly dependent on z in the first and third bins, but there is little dependence in the second bin. Looking at the right panel, we see that misspecification of the cutoffs, however, induces a nonrandomization 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-bin dependence does not exist in the true data-generating process.

As we noted in Section 4, we handle this in practice by including z directly in each bin-level model. In particular, we now 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 A.3 presents the analogous performance plots for this situation across 100 replications. Panel (a) in Figure A.3 presents the case with nonrandom within-bin allocation and correctly specified cutoffs, panel (b) in Figure A.3 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 nonrandomness 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 A.3. Performance of Estimator with Nonrandom Within-Bin Allocation and Misspecified Cutoffs



Discussion. We conclude this appendix with a brief discussion of some limitations of the procedure that may be useful in determining its use in other settings.

First, there is loss in efficiency from not pooling across segments. This is likely to be a concern in sparse data applications. On the other hand, in database applications in which there is a large number of observations within each segment, this concern may not be practically important. Second, 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 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 the proposed method. More generally, when choosing a model for heterogeneity, a researcher always faces the usual trade-off of utilizing the "within" versus 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 trade-offs between these two considerations. This assessment applies to the decision to use the discussed method as well.

Finally, it is important to note that this 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 with 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 setup will therefore need to be augmented with a model of how *theo* evolves in response to promotions. Then, Equation (7), which specifies the process for z, will have to be fully specified. We avoided modeling this because of model complexity and informational constraints, as our goal was to present a way to do estimation without needing to specify the full likelihood. Yet doing this may become important if the goals become more ambitious and assessments of long-run effects are required.

Appendix B. Tables and Figures

Table B.1. Segment Definitions Used in the Analysis

Segment num.	Segment bin	Proportion of consumers (%)	No. 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

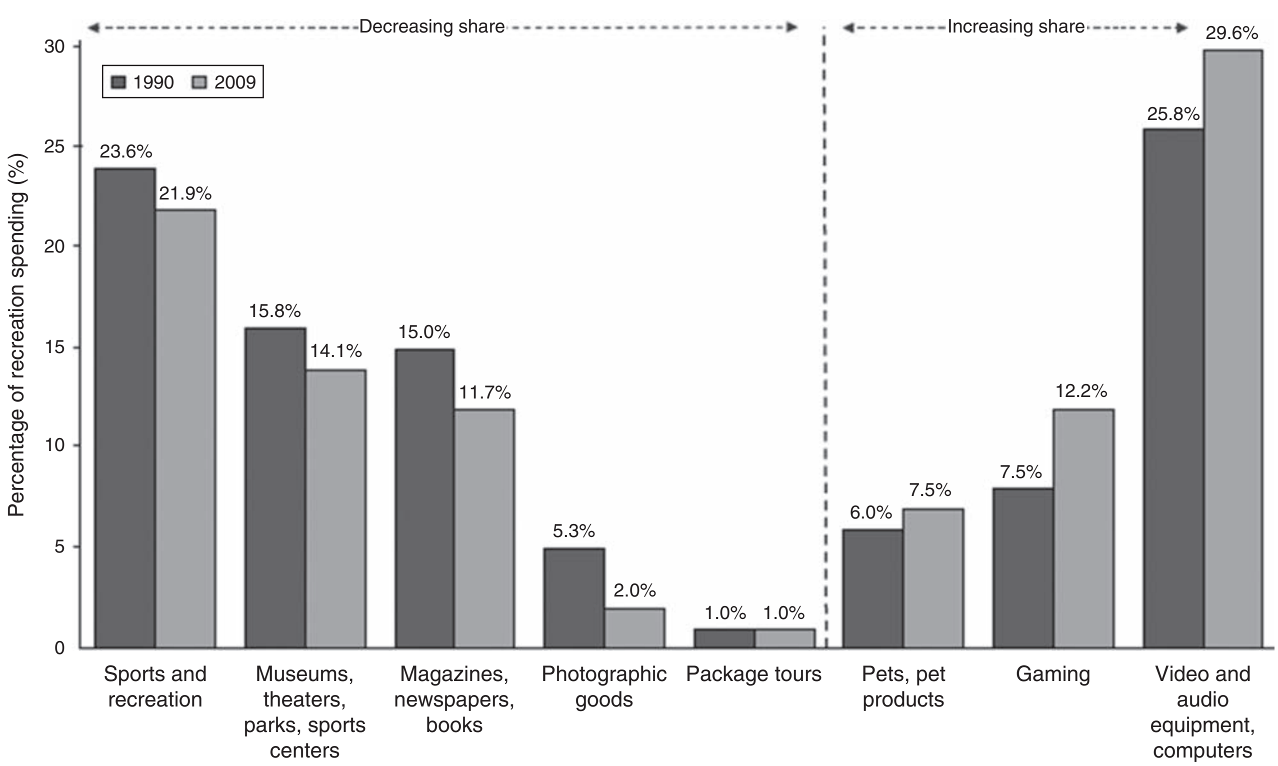
Appendix B. Tables and Figures (Continued)

Table B.1. (Continued)

Segment num.	Segment bin	Proportion of consumers (%)	No. of trips
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

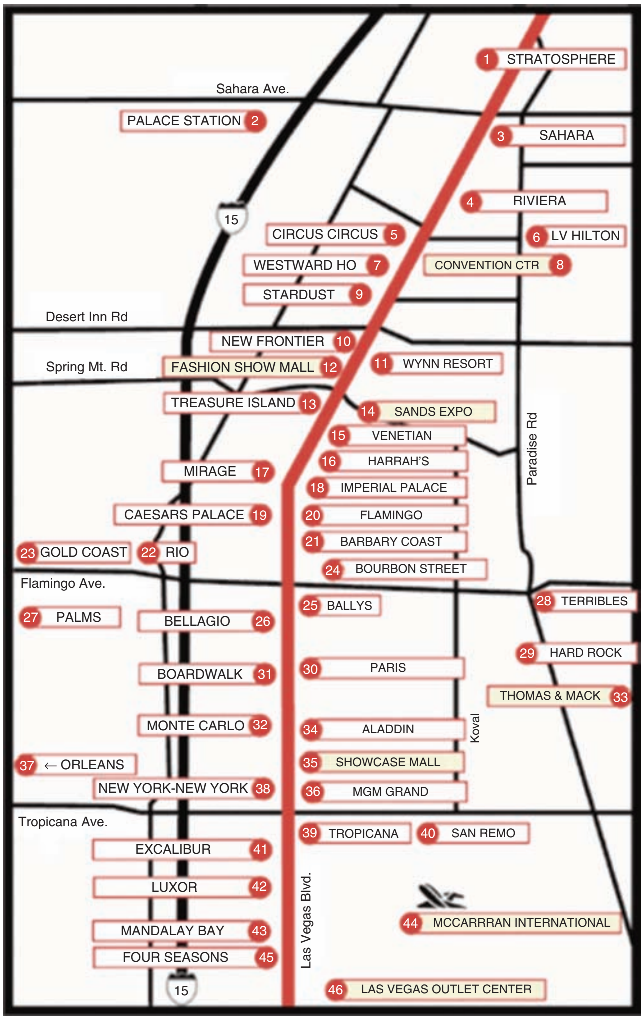
Notes. Segments are based on the number of trips observed, range of *theo*, and location and slot/table preference. Data span January 2008 to July 2010. For confidentiality reasons, an undisclosed, random number of bins are omitted from the table (so the percentages will not add up to 1).

Figure B.1. Consumer Spending on Recreation Between 1990 and 2009



Source. Bazelon et al. (2012).
Notes. Numbers are based on total recreation expenditure of $314.7 billion in 1990 and $897.1 billion in 2009. "Video and audio equipment, computers" includes video and audio equipment, information processing equipment, and services related to video and audio goods and computers. "Sports and recreation" includes sports and recreational vehicles, other sporting and recreational goods, and maintenance and repair of recreational vehicles and sports equipment. "Museums, theaters, parks, sports centers" includes membership clubs and participating sports centers; amusement parks, campgrounds, and related recreational services; admissions to specified spectator amusements; motion picture theaters; live entertainment, excluding sports; spectator sports; and museums and libraries.

Figure B.2. (Color online) Agglomeration of Commercial Casinos on the Las Vegas Strip



Endnotes

[1] This estimate is based on commercial casino revenues reported in American Gaming Association (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.

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

[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.

[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.

[5] 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.

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

[7] 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 is not obtained 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.

[8] An early working version of this 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 their interactions in the model for y. The Monte Carlo simulations reported in Appendix A assess sensitivity to this issue.

[9] 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 subpopulation of agents who self-selected into 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.

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

[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 out this kind of promotion adjustment at the individual property level, we believe the impact of these on our results is small, as these kinds of activities account for less than 5% of the promotional spending by the properties. The bulk of the property-specific promotions were mailed out and were predetermined during the intervention period.

[12] The economic cost involves two components: (1) the money that could have been made if the promoted product were offered to the next best customer instead of the focal customer and (2) the counterfactual profit that could have been made from the focal customer in the absence of the promotion being offered to him. To compute the first part, we would have to assess whether a promoted product had binding capacity constraints and could have been sold (at potentially different prices) to other consumers in current or later periods, and then assess the profit implication of that strategy. While in some cases this can be evaluated (or at least approximated) with sophisticated yield management systems, MGM did not have these fully built out in an accessible way for all capacity-constrained offerings (hotel rooms, restaurants, shows, etc.) The multidimensional nature of the promotion bundles on offer increases the complexity of assessing such costs, so computing this ourselves would involve additional assumptions, potentially introducing a significant element of subjectivity into the process. On the second component, the outcomes in the control group represent the right counterfactual for what would happen to consumer behavior in the absence of receiving corporate promotions. In particular, consumers in the control group can choose to utilize the individual property-specific promotions they could substitute to when corporate promotions are not offered. The outcomes in the control represent the revenue/costs of these, so the comparison to the control accommodates this aspect of costs when assessing treatment effects. We thank the review team for helpful comments on this issue.

[13] As mentioned in Section 4.1, campaign planning in the status quo reflects various objectives (like increasing visitation), while the model/academic allocation reflects profit maximization as the objective. Though possible, the contrast between the model/academic and

the status quo allocations is unlikely to represent in a significant way the differences in these objectives. The campaign-specific objectives of the firm are mainly reflected in the choice of the promotions to include in the superset of offers in the campaign: this set is held the same in the comparison. Once the promotion bundles are picked, the corporate marketing teams implementing the new allocation and the status quo allocation face the same incentive—to allocate those picked bundles to the right set of customers in the way that will produce maximal return. Our understanding is this is what the status quo represents. In the event that this is not true, the treatment effect should be interpreted as reflecting the effect of the new allocation plus that of the profit-oriented objective of the model/academic–based policy.

[14] A related question here is why the redemption costs in Groups A and B are less that the scaled costs of the control group, even though both Groups A and B are also exposed to similar property-level promotions as the control. The reason is that consumers in Groups A and B 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 A (respectively, Group B) 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 to clarify this. Suppose a household in Group A 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.

[15] Compared to no corporate marketing (but continuing the individual property-specific promotions), the new policy produces a net incremental ROI of 0.22.

References

American Gaming Association (2012) State of the states: The AGA survey of casino entertainment. https://ok/research/reports/state-states-12.

Ansari A, Mela C (2003) E-customization. *J. Marketing Res.* 40(2):131–145.

Arora N, Dreze X, Ghose A, Hess JD, Iyengar R, Jing B, Joshi Y, et al. (2008) Putting one-to-one marketing to work: Personalization, customization, and choice. *Marketing Lett.* 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, Washington, DC.

Bucklin RE, Gupta S (1999) Commercial use of UPC scanner data: Industry and academic perspectives. *Marketing Sci.* 18(3):247–273.

Busse M, Simester D, Zettelmeyer F (2010) The best price you'll ever get: The 2005 employee discount pricing promotions in the U.S. automobile industry. *Marketing Sci.* 29(2):268–290.

Chiang J (1995) Competing coupon promotions and category sales. *Marketing Sci.* 14(1):105–122.

Cho S, Rust J (2008) Is econometrics useful for private policy making? A case study of replacement policy at an auto rental company. *J. Econometrics* 145:243–257.

Encyclopedia.com (2012) Commercial casinos. http://www.encyclopedia.com/topic/Casino.aspx.

Goldfarb A, Tucker CE (2011) Online display advertising: Targeting and obtrusiveness. *Marketing Sci.* 30(3):389–404.

Hartmann W, Nair H, Narayanan S (2011) Identifying causal marketing mix effects using a regression discontinuity design. *Marketing Sci.* 30(6):1079–1097.

Leeflang P, Wittink D (2000) Building models for marketing decisions: Past, present and future. *Internat. J. Res. Marketing* 17(2–3):105–126.

Lilien G, Roberts J, Shankar V (2013) Effective marketing science applications: Insights from the ISMS-MSI Practice Prize finalist papers and projects. *Marketing Sci.* 32(2):229–245.

Little JD (1970) Models and managers: The concept of a decision calculus. *Management Sci.* 16(8):B-466–B-485.

Lodish L (2001) Building marketing models that make money. *Interfaces* 31(3):S45–S55.

Manchanda P, Rossi PE, Chintagunta PK (2004) Response modeling with non-random marketing mix variables. *J. Marketing Res.* 41(4):467–478.

Mantrala M, Seetharaman PB, Kaul R, Gopalakrishna S, Stam A (2006) Optimal pricing strategies for an automotive aftermarket retailer. *J. Marketing Res.* 43(4):588–604.

Misra S, Nair H (2011) A structural model of sales-force compensation dynamics: Estimation and field implementation. *Quant. Marketing Econom.* 9(3):211–225.

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

Roberts J (2000) The intersection of modeling potential and practice. *Internat. J. Res. Marketing* 17(2–3):127–134.

Rossi PE, McCulloch RE, Allenby GM (1996) The value of purchase history data in target marketing. *Marketing Sci.* 15(4):321–340.

Sahni N, Zou D, Chintagunta P (2017) Do targeted discount offers serve as advertising? Evidence from 70 field experiments. *Management Sci.* 63(8):2688–2705.

Simester D, Sun P, Tsitsiklis J (2006) Dynamic catalog mailing policies. *Management Sci.* 52(5):683–696.

Simester D, Hu J, Brynjolfsson E, Anderson E (2009) Dynamics of retail advertising: Evidence from a field experiment. *Econom. Inquiry* 47(3):482–499.

Sinha P, Zoltners A (2001) Sales-force decision models: Insights from 25 years of implementation. *Interfaces* 31(3):S8–S44.

Sudhir K, Roy S, Cherian M (2016) Do sympathy biases affect charitable giving? The effects of advertising content. *Marketing Sci.* 35(6):849–869.

Van Bruggen GH, Wierenga B (2001) Matching management support systems and managerial problem-solving modes: The key to effective decision support. *Eur. Management J.* 19(3):228–238.

Winer RS (2000) Comments on Leeflang and Wittink. *Internat. J. Res. 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, University of Pennsylvania, Philadelphia.