An Empirical Analysis of Individual Level Casino Gambling Behavior

Sridhar Narayanan
Puneet Manchanda*

June 2006

PRELIMINARY AND INCOMPLETE: PLEASE DO NOT CIRCULATE WITHOUT PERMISSION

*Sridhar Narayanan is Assistant Professor of Marketing at the Graduate School of Business, Stanford University and Puneet Manchanda is Associate Professor of Marketing at the Graduate School of Business, University of Chicago. The authors are grateful to Wes Hartman, Harikesh Nair and the participants of the 2006 Marketing Science conference for comments and feedback and an anonymous company for providing the data and institutional insights. The second author would also like to acknowledge the Kilts Center for Marketing at the University of Chicago for providing research funds and the True North Faculty Fund at the Graduate School of Business for research support. All correspondence may be addressed to the authors via e-mail at *sridhar.narayanan@gsb.stanford.edu*(Narayanan) or *puneet.manchanda@ChicagoGSB.edu* (Manchanda).

Abstract

Gambling and gaming has evolved to becoming a very large and pervasive industry in the Unites States over the last three decades. The industry is worth over $ 73 billion in revenues. Participation in this industry is around one-third of all adult Americans. The nature of this industry and its rapid growth has led to a lot of debate about its benefits and costs. In this paper,our access to a rich and new dataset on individual consumer behavior vis-a-vis casino visitation and activity allows us to take a data based approach to investigating some of the commonly made arguments about these benefits and costs.

We focus our attention on one of the commonly cited costs of gambling i.e., that it leads to addictive behavior (with potentially harmful individual and societal effects). We use the commonly accepted definition of addiction from the economics literature to test for its presence i.e., that current consumption is affected by past consumption. We fit a model of the play decision and bet amount to data from a consumer panel of casino visitors over a two year period. Our data are at a highly disaggregate level – we look at play decisions within a given trip for individual consumers. Our modeling approach allows us to exploit the rich variation in the data both across and within individuals.

Our results show that, controlling for other reasons that could induce play, only about 7% of all consumers show evidence for addiction. While this proportion may look small, it is consistent with research carried out in other domains that has focussed on casino gamblers. In addition, we find no evidence for either the hot hand myth or the gambler's fallacy. An interesting (though indirect) implication of our analysis is that casino gambling seems to be another form of entertainment for the non-addicted consumers.

Our paper can also be seen as more descriptive in terms of the role of marketing in this industry. We find that marketing activity has a positive effect on the decision to play and the amount to bet. In terms of effect size, comps seem to be more similar to advertising rather than price promotions. Finally, we find some weak evidence that marketing activity is more effective for consumers who exhibit more addictive behavior.

Keywords: *Gambling Behavior, Addiction, Selection Models, Hierarchical Bayes Methods*

1 Introduction

Gambling and gaming has evolved to becoming a very large and pervasive industry in the Unites States over the last three decades. In terms of size, it is estimated to have a total revenue of $ 73 billion in 2003. In terms of participation, a 2004 Gallup Poll showed that about half of all adult Americans bought a state lottery ticket and about one-third of all adult Americans visited a casino in 2003. The size and pervasiveness of the industry has led to extensive debate about the benefits and costs of this industry to society. Supporters of the industry point to the fact that it provides an entertainment outlet for consumers, facilitates job creation, provides additional tax revenues and increases the welfare of disadvantaged groups in society such as Native Americans. Opponents of the industry argue that the rapid expansion of the industry creates opportunities that lead to undesirable behavior modification such as "addiction" and increased propensity for criminal behavior. The current academic literature also does not provide any definitive findings on the role of gambling at the individual consumer level.

The major objective of our study is to shed light on the role of gambling at an individual customer level. Using a rich, micro-level panel data of consumer betting activity at a casino, we try to find evidence for addiction while controlling for win/loss outcomes and marketing activity directed at the consumer. The panel nature of our data allows us to test for this at the individual customer level as we can obtain individual consumer level estimates. This is crucial as previous (medical) studies on gambling addiction have found that addicitive behavior is detected in only a small proportion of consumers (Potenza, Kosten, and Rounsaville 2001). Interestingly, if we are unable to find evidence for addiction, the logical conclusion that can be drawn is that casinos do provide entertainment. An interesting facet of our analysis is that we can investigate the effectiveness of marketing activity in this industry and its interaction with addictive behavior (if it is indeed present).

As the main objective of this paper is to test for addictive behavior in casino gambling, we need to define what we mean by addiction. We use the well accepted economic definition of addiction (Pollak 1970, Becker and Murphy 1988). This literature defines addiction as a positive effect of past consumption on the marginal utility of current consumption. In the reduced form, this definition implies that there is a positive correlation in consumption across time periods. In our scenario, we expect to find this positive correlation in consumption in a consumer's decision to play and bet an amount as a function of his/her last play. Other factors that could lead to correlation in play and bet amount across occasions could be the immediate past wins (the "hot hand myth") or the immediate past losses (the "gambler's fallacy"), the cumulative time spent in the casino and the effect of marketing.

We therefore specify a joint model of a consumer's decision to play (given that s/he has already played) and given the play decision, the amount bet as a function of past amount bet and the control variables described above. We cast this model in a hierarchical Bayesian framework and estimate the model using Markov chain Monte Carlo methods. The advantage of using this approach is that we are able to obtain individual level parameters that help inform about the role of gambling for individual consumers.

Our results suggest that there is not much evidence for addiction in gambling in the casino our data is from, though there is a small subset of consumers who show some degree of addictive behavior. We also do not find evidence for consumers' belief in either the hot hand myth or in the gambler's fallacy. Consumers respond to their overall level of wins or losses, and are more likely to participate and bet a higher amount if they have been winning than if they have been losing money during a trip. Marketing activities by the casino significantly affect both consumers' play decisions and the amounts they bet. Consumers who display evidence for addictive behavior are more likely to have higher average bet amounts and are also more likely to be affected by marketing activities.

The rest of the paper is organized as follows. We briefly discuss the relevant literature in §2. §3 describes the data. We describe our model and the econometric specification in §4 and present some details of our estimation in §5. We discuss the results in §6 and conclude in §7.

2 Literature Review

We provide a very brief introduction to the literature. We draw on the literature from two areas - the economics literature on addiction and the behavioral literature on gambling. Early studies in the economics of addiction literature (Pollak 1970) modeled addiction through the dependence of past consumption on current consumption. Subsequent work (Stigler and Becker 1977, Becker and Murphy 1988) developed the idea of rational addiction, where consumers are forward-looking and therefore account for the effect of current consumption on future utility when making their consumption decisions. There is also a long empirical literature on addiction, that has studied addictive behavior in a variety of contexts. Some examples of the contexts where addiction has been studied are cigarette smoking (Chaloupka 1991, Becker, Grossman, and Murphy 1994, Gruber and Koszegi 2001), caffeine (Olekalns and Bardsley 1996) and gambling in state lotteries (Guryan and Kearney 2005). We draw from this literature in developing our model and use the definition of addiction in this literature.

At this point, it should be clarified that we do not make a distinction between habit formation and addiction, both of which are forms of state dependence. Indeed, the literature on addiction itself has often used these terms somewhat interchangeably (Pollak 1970). Further, studies both in economics (Heckman 1981) and in marketing (Roy, Chintagunta, and Haldar 1996) have used habit persistence to refer to dependence of current choice on previous *propensity* to be in a state rather than the actual state itself and have distinguished this from structural state dependence, which is the dependence on the previous state. In our context, we are modeling the latter effect rather than the former. Overall, we are agnostic about whether this effect is termed addiction or habit formation and will use the term addiction henceforth to describe it.

There is a vast literature on gambling across a variety of domains. We draw on this literature to inform our modeling choices and test for two common behaviors that have been documented in this domain. The first behavior is usually labelled as the "hot hand myth." This refers to the mistaken belief that two consecutive draws from a random distribution are positively correlated. Besides the long tradition of studies based on laboratory experiments that documents this behavior, empirical studies of state lotteries (Guryan and Kearney 2005) also find evidence for it. The second behavior, commonly referred to as the gambler's fallacy, is a belief that two consecutive draws are negatively correlated. This has also been documented, again in the context of state lotteries (Clotfelter and Cook 1993). It is not clear *a priori* which, if any, of these effects may be present in the case of casino gambling and hence we allow for our modeling approach to accomodate either explanation.

In sum, while a lot is known about addiction and consumers' propensity and biases vis-a-vis gambling, there is little empirical verification of these effects using individual behavioral data. Our endeavor in this paper is to complement the existing literature in these two domains.

3 Industry Background and Data

As mentioned earlier, the gambling industry is estimated to have a total revenue of $ 73 billion in 2003. The industry as a whole can be divided into three major sub-categories. The largest sub-category is casino gaming accounting for 54% of this revenue. The next largest sub-category is that of state lotteries (25%). The balance is accounted for by a variety of smaller activities such as legal bookmaking and charitable games. In terms of consumer behavior, about half of all adult Americans buy a lottery ticket and about one-third of all adult Americans visit a casino every year (Gallup Poll 2003).[1] These numbers are confirmed by surveys conducted by casino operators e.g., a study carried out by Harrah's finds that one-fourth of all adult Americans has visited a casino in the last one year (Jones 2004). They also find that the average visit frequency is 5.8 times a year. Interestingly, the anecdotal belief in the industry is that consumer participation has "saturated". This is confirmed by the same Harrah's study that the number of trips to casinos in the aggregate has grown at a negligible pace over the last four years. This has resulted in many casino companies trying to increase "share-of-wallet" and has increased the attention given to the role of marketing in the industry.

Our data come from a leading casino company operating primarily in the United States. The company operates properties that are spread nationwide (i.e., not just concentrated in Las Vegas in Nevada).

The company operates a loyalty card program to facilitate the building of relationships with its consumers. Enrollment in the loyalty card program is free and most consumers sign up for the card as it facilitates communication between the consumer and the company. Membership in the loyalty program is ubiquitous – the company estimates that cardholders account for about 80% of all play in its properties. Typically, the consumer swipes the card

[1] The details of this poll, including the number of adult Americans who buy a lottery ticket are referred to in Kearney (2005)

before s/he begins play. All activity is then uniquely linked to that customer's account.

Our data come from one property owned by the company.[2] This property is located in the southwestern United States. The property is a "local monopoly" as it is located about 30 miles from the nearest competitor casino (about 45 minutes driving time) and at an average distance of about 60 miles from other casinos in that part of the state.

The property is open for operation twenty-four hours a day and seven days a week and offers play via slot machines, video poker, keno, bingo, blackjack and craps. The majority of amount bet - 90% - in this property comes from slots.

Customer play data is available to us for a two year period (Jan 2004-Dec 2005). The play data comprise all bet amounts by each consumer on each visit made to the casino. Besides the bet, we also observe the win/loss on each play, the comp amount delivered to the customer on a daily basis and some demographics of each consumer (age, gender, race).

Our approach is to use the lowest level of aggregation in a manner that we can best isolate the effects of past consumption on current consumption. We therefore treat the play within a customer trip as our unit of analysis. A play is initiated when a customer swipes his/her loyalty card at a station (typically a slot machine) and begins betting. The play ends when the customer exits the station. A trip is defined a sequence of temporally contiguous plays (over either a single day or a sequence of contiguous days). Our assumption therefore is that there is no cross-trip dependence.[3] Thus, all the predictors of the play and bet activity used by us get reset at the beginning of a trip.

The company markets its properties and games to consumers via coupons (redeemable for free hotel stays and/or meal) or cash equivalents. Consistent with industry practive, the company divides the use of these instruments into two broad classes. The first class is referred to as "offers." Offers essentially consist of all marketing activity that is aimed at influencing the decision to visit a given property (or set of properties). The second class is referred to as "comps." Comps are geared towards increasing the duration of play and the amount bet once a consumer has shown up at the property and has begun play. The

[2]Due to confidentiality concerns, we cannot reveal the name of the company and the geographic location of the property.

[3]This could be an unreasonable assumption if there is a large proportion of trips for a given consumer that are separated only by a few days.

company records the dollar value of comps delivered to each customer on each trip.

The total number of consumers in our data is 198,223. These consumers comprise the universe of all players with loyalty cards who are affiliated with this property in this time period. For the purpose of our analysis, we take a random sample of 2000 consumers from this set. The total number of trips made by these consumers was 15632 with the number of mean trips per consumer being 7.8 (with a standard deviation of 17). This averages out to about 4 trips per year which is close to the national average (5.8) described above. In a trip, these consumers play on average 3.5 times (standard deviation 2.1) resulting in 56142 bets in total.

We provide descriptive statistics for trip level activity in Table 1. As can be seen from the table, the median trip lasts two hours, the total amount bet during a trip is $ 450.00 (3.5 bets times $ 131.00) . The median total cumulative win is $ -46. In other words, the casino holds back about $ 46 out of a total amount bet of $ 450.00 i.e., a gross margin of about 10% (which is consistent with the anecdotal evidence on margins of slot machines). For our purpose, the relevant marketing activity are comps - the median comp value amount is 0. The mean comp amount is $ 9 and is two percent of the amount bet. Given that the gross margin is 10%, this suggests that the firm spends 20% of its gross revenues on comps. Note also that the data are somewhat skewed to the right (the contrast between the median and the mean is quite high).

In terms of the demographics, the proportion of males is 45.9%, the mean age is 61 (standard deviation 13) and the racial/ethnic composition is as follows - Western European (64.6%), Hispanic (10%), Mediterranean (4.6%), Eastern European (3.7%) and Other (17.1%). Thus, the typical consumer in our sample is an older, white female.

Table 1 about here

4 Model

4.1 Model Development

We now discuss our model development. We describe the three important aspects of our model - addiction, selection and heterogeneity - below.

As mentioned earlier, we adopt the commonly accepted reduced form specification for addiction i.e., we allow for dependence in play and consumption across plays (bets) for a consumer within a trip to the casino. In our current formulation, we assume that consumers are myopic. In other words, current consumption is affected only by the (immediate) past consumption. This is in contrast to models of rational addiction in which all consumption - past and expected future consumption - has an effect on current consumption. Our rationale for doing this is twofold. First, some previous research on gambling has not found evidence that future consumption affects current consumption (Guryan and Kearney 2005). Second, our context makes it hard to model the effect of future consumption on current consumption. This is because the decision whether to play again within a trip after the current bet cannot be assumed to be exogenous (see the discussion below). Since the expected value of a bet is negative, greater consumption in the current period would lead to greater expected losses *ex ante*, which in turn would reduce the probability of participation in a future bet. Thus, the effect of endogenous participation is opposite to that of forward looking behavior in this context. Modeling this complex set of factors is challenging and we defer this to future research.

An important consideration when developing our model is that we only observe the amount of a bet placed by a consumer when she chooses to participate in the bet. The decision to participate in a bet is not independent of the amount of the bet. This leads to a classic selection problem. Specifically, if there are unobserved factors that affect both the decision whether to participate in a bet as well as the amount of the bet, the estimates from a regression of the bet amount would be biased. Hence, we need to control for this selection. Heckman (1979) proposed a two-stage estimator for correcting for this selection bias. Given our Bayesian model, we cannot directly use this two-stage approach to control for selection. Hence, we develop an approach where we directly model the correlation between the bet amount and participation decisions.

Given this structure, for identification purposes, we need an instrument that affects the participation decision, but not the amount of the bet. The instrument we use is the length of time that the customer had played in the casino on a given trip. As time spent on a trip increases, the consumer has to make a decision whether to keep playing or not. The amount of time spent would be a proxy for satiation and/or fatigue and thus directly influence the decision to play. On the other hand, given that the consumer is going to play, it is unlikely that the amount of time spent playing so far will affect the bet amount during the next play, controlling for the total amount of money won or lost until then.

Finally, we allow for heterogeneous response on all our covariates. Our hierarchial Bayesian approach allows us to obtain the posterior of distribution of population as well as individual level parameters. Obtaining the individual level parameters is crucial as prior research has documented that only a small proportion of consumers behave in an addictive manner while gambling (Potenza, Kosten, and Rounsaville 2001). Therefore, testing for the presence of addictive behavior for the average consumer may lead to a conclusion that this effect does not exist.

4.2 Model Specification

The model consists of two sub-models, one sub-model for the bet amount and another for the decision of whether to participate in the bet at all. We shall henceforth refer to the two as the *bet* sub-model and the *play* sub-model respectively.

The bet sub-model is given as

$$\ln(Bet_{it}) = X_{it}\alpha_i + \varepsilon_{it}, \text{ if } Play_{it} = 1 \tag{1}$$

where Bet_{it} is the amount of the bet in dollars, $Play_{it}$ is an indicator variable for whether the consumer decides to participate in the bet or not, X_{it} is a vector of observable covariates, α_i is an individual-specific parameter vector and ε_{it} is an unobservable factor that affects the bet amount.

The decision whether to play or not is assumed to be driven by an underlying latent variable that we term $Play^*_{it}$ given by the following equation

$$Play_{it}^* = Y_{it}\beta_i + \nu_{it} \tag{2}$$

subject to the following constraints

$$\begin{array}{ll} Play_{it}^* > 0 & \text{if } Play_{it} = 1 \\ Play_{it}^* < 0 & \text{if } Play_{it} = 0 \end{array} \tag{3}$$

Thus, (2) and (3) jointly describe the play sub-model.

As described earlier in section 4.1, we need to control for selection by modeling the correlation between the bet equation and play equation. Thus, we make the assumption that ε_{it} and ν_{it} have a joint bivariate normal distribution with a non-diagonal covariance matrix.

$$\begin{pmatrix} \varepsilon_{it} \\ \nu_{it} \end{pmatrix} \sim N\left(0, \begin{pmatrix} \sigma^2 & \rho\sigma \\ \rho\sigma & 1 \end{pmatrix}\right) \tag{4}$$

Here, the correlation between the bet equation and the play equation is ρ. Note also that the variance of the marginal distribution of the play equation is fixed to 1 for identification purposes.

Thus, the bet sub-model is a log-linear regression and the play sub-model is a probit model. The correlation ρ explicitly controls for selection and we have an excluded variable, the length of time spent in the casino until t in the covariates for the play sub-model Y_{it} but not in the covariates for the bet sub-model X_{it}. Note also that all the variables that are in X_{it} can be in Y_{it}.

Combining 1, 2, 3 and 4, we can write the joint model as

$$\begin{array}{ll} \begin{pmatrix} \ln(Bet_{it}) \\ Play_{it}^* \end{pmatrix} \sim N\left(\begin{pmatrix} X_{it}\alpha_i \\ Y_{it}\beta_i \end{pmatrix}, \begin{pmatrix} \sigma^2 & \rho\sigma \\ \rho\sigma & 1 \end{pmatrix}\right) 1\left(Play_{it}^* > 0\right) & \text{if } Play_{it} = 1 \\ \textit{and} & \\ Play_{it}^* \sim N\left(Y_{it}\beta_i, 1\right) 1\left(Play_{it}^* < 0\right) & \text{if } Play_{it} = 0 \end{array} \tag{5}$$

Thus, if the consumer decides to play, the model is a truncated bivariate normal model, and a truncated univariate normal model otherwise.

4.3 Hierarchical Model

Since the model is specified at the individual consumer level, we need to specify the hierarchy for the individual-level parameters. We assume that the individual level parameters in the bet and play sub-models are a function of observed demographic variables z_i and are distributed in the population as follows

$$\alpha_i \sim N\left(M_\alpha z_i, V_\alpha\right) \tag{6}$$

$$\beta_i \sim N\left(M_\beta z_i, V_\beta\right) \tag{7}$$

The model is completed by specifying the prior distributions for all the population-level parameters (σ^2,ρ,M_α, V_α, M_β, V_β).

$$\sigma^2 \sim \frac{\nu_\sigma s_\sigma^2}{\chi^2_{\nu_\sigma}} \tag{8}$$

$$\rho' = \frac{\rho + 1}{2} \sim \text{Beta}\left(\lambda_1, \lambda_2\right) \tag{9}$$

$$\text{vec}\left(M_\alpha\right)|V_\alpha \sim N\left(\bar{\alpha}, V_\alpha \otimes A^{-1}\right) \tag{10}$$

$$V_\alpha \sim \text{Inverse Wishart}\left(\mu_\alpha, S_\alpha\right) \tag{11}$$

$$\text{vec}\left(M_\beta\right)|V_\beta \sim N\left(\bar{\beta}, V_\beta \otimes B^{-1}\right) \tag{12}$$

$$V_\beta \sim \text{Inverse Wishart}\left(\mu_\beta, S_\beta\right) \tag{13}$$

5 Estimation

As mentioned in the data section, we estimate our model on a random sample of 2000 customers. For the sake of estimation, we lose one observation per trip, since we include the previous bet amount in the specification to control for addiction. This leaves us with 40510 observations, although we have 56142 observations for the play sub-model (since there is one last observation of the consumer not choosing to play). Within a trip, the vector of $Play_{it}$ looks like the following

$$Play_{it} = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & ... & 0 \end{pmatrix}' \tag{14}$$

The covariates for the bet sub-model (X_{it}) include, besides an intercept, the bet amount in the previous bet within the trip, the amount won or lost in the previous bet, the cumulative amount won or lost in the trip until time period t and the amount of comps the consumer has received. The coefficient on previous bet is our "test" for addiction, the coefficient on amount won or lost in the previous bet tests for the hot hand myth or gambler's fallacy, the total amount won or lost controls allows us to exclude the length of time spent in the casino in the play equation and comps control for marketing activities. Thus,

$$X_{it} = \begin{pmatrix} 1 & LastBet_{it} & LastWin_{it} & CumWin_{it} & Comps_{it} \end{pmatrix} \tag{15}$$

The covariates for the play sub-model Y_{it}) include all the covariates in the bet sub-model (X_{it}) and additionally, the excluded variable, which is the time spent in the casino until t. Thus,

$$Y_{it} = \begin{pmatrix} 1 & LastBet_{it} & LastWin_{it} & CumWin_{it} & Comps_{it} & TimeSpent_{it} \end{pmatrix} \tag{16}$$

Currently, we do not include demographic variables in the specification for the individual-level parameters α_i and β_i. Hence, z_i is simply an identity vector and M_α and M_β are vectors rather than matrices.

The estimation of this model requires us to draw from the joint posterior distribution of all parameters conditional on the data. Since the joint distribution is not from a known family of distributions and cannot be drawn from directly, we use the Gibbs Sampler (Gelfand and Smith 1990), making sequential draws from the full conditional distributions of sub-vectors of the parameter vector. At convergence, these sequential draws from the full conditional distributions give us draws from the joint distribution of the parameter vector. Further, we use data augmentation (Tanner and Wong 1987) and draw from the full-conditional distribution of $Play^*_{it}$, taking into account the fact that it is a truncated normal distribution.

The derivations of the full conditional distributions of the various parameters are given in the Appendix. Note that the full conditional distributions of σ^2 and ρ are not from known distributional families. Hence, we use the Random Walk Metropolis-Hastings algorithm (Chib and Greenberg 1995) to make draws from the full conditional distributions of these parameters. The remaining sub-vectors of the parameter vector are drawn directly since they are from known distribution families.

The chain was run for a total of 30,000 draws, with the first 25,000 draws being discarded and the inference done on the remaining 5,000 draws. Convergence was assessed by inspecting the chains for each of the parameters.

6 Results

We now discuss our results. We first describe the population level parameters and then we examine the distributions and correlations across the individual level parameters.

6.1 Population Level Parameters

We report the population level parameters in table 2. We find that most of the population-level parameters are significantly different from zero. The coefficient for $LastBet$ is negative at the population-level in both the bet sub-model and the play sub-model. This suggests that, all else being equal, consumers are on average less likely to play again if they had a higher previous bet, and conditional on participation, their bet amount is also lower. This is evidence against addiction in gambling behavior, at least at the population average level.

The next parameter of interest is the coefficient on $LastWin$. This coefficient is a test of the typical fallacies based on the "law of small numbers" - the hot hand myth and the gambler's fallacy. The hot hand myth refers to the mistaken belief that a previous win is positively correlated with the likelihood of a future win. The gambler's fallacy, on the other hand, is the opposite of this, and refers to the belief that the probability of a win in an independent future game is negatively correlated with a past win. If the coefficient on $LastWin$ is positive, it would provide us evidence for the hot hand myth, while a negative coefficient would provide evidence for the gambler's fallacy. We find that this parameter is not significantly different from zero. This suggests that in at least the casino gambling, there is not much evidence for either the hot hand myth or the gambler's fallacy.

The coefficient for $CumWin$ is positive and significantly different from zero in both the bet and play sub-models. This variable represents the total amount of money that the consumer has lost or won during the entire trip and proxies for the wealth effect on gambling behavior. A positive coefficient in the play sub-model is intuitive, since we are more likely to see continued participation from consumers at a particular point of time if they have won money until then and lower participation as a result of consumers reaching a threshold level of losses (or losing all their money, at the limit). On the other hand, the prediction for the effect on bet amount is more ambiguous, since consumers might on the one hand be induced to bet more if they win more, but also might want to cash in their wins. Thus, it is interesting that we find a positive effect of the total amount of wins on the bet amount.

The effect of *Comps* is positive and significant for both the bet and play sub-models. Thus, it appears that marketing activity is positively related to both the bet amount and the probability of participation. Finally, in the play sub-model, the excluded variable, which is time of play, has a negative coefficient. As discussed before, this may be due to satiation, play fatigue or time constraints in general.

Finally, the correlation between the play and bet equations, represented by ρ, is positive but relatively small at 0.25. This confirms our expectation that selection plays a role in consumer behavior on casino trips.

Table 2 about here

6.2 Individual Level Parameters

Next we look at the individual-level parameters, which are reported in Table 3.[4] Recall that it is crucial for us to be able to estimate the individual parameters with a reasonable degree of certainty so that we can test for the various effects that we have outlined. Our approach allows us to compute the posterior distribution of each parameter for each individual. A natural question to ask is if the posterior distribution of these parameters are tightly estimated. We use the ratio of the posterior mean to the posterior standard deviation as our informal metric to determine is each individual parameter is "significant." Our cutoff to determine this significance is 1.65 (akin to the t-statistic at p=0.10). Across both the play and the bet equation, a large proportion of these individual-level parameters are significant. The intercept terms for both equations are significant for over 95% of the households. The coefficients for $Comps$ are also significant for over 90% of the households in both equations. The coefficients for $LastBet$ are significant for about 70% of the households. The individual level coefficients for $LastWin$ are significant only for about 10% of the consumers and that for $CumWin$ for about 30% of the consumers.

The parameters for addiction are the coefficients for $LastBet$. If this parameter is positive, it suggests positive dependence of consumption over time and indicates evidence for addictive behavior. In the bet sub-model, it is significantly positive only for a total of 10 out of 1769 consumers for which we have this parameter estimate. The coefficient is either significantly negative or not significant for the remaining consumers. In the play sub-model, it is positive for a total of 143 consumers i.e., 8% of our sample. We then examine how tightly the parameter is estimated for these consumers. We find that it is significant for 124 or 87% of consumers with a positive coefficient. In contrast, for the remaining consumers i.e. for whom we have a negative coefficient, this is significant for 69%. Thus, we find evidence for addictive behavior for a relative small group (7%) of consumers. It is interesting to note that this proportion of consumers who display some evidence of addictive behavior in terms of the effect of past bet amount on current probability of participation is comparable to

[4] We report the individual-level parameter estimates for 1769 out of 2000 consumers in the estimation sample. For the remaining 231 consumers, the parameters of the bet equation are not defined since these are households for which there were no trips with more than one play. We have play equation parameters for these households, but to ensure consistency, we report play equation parameters also for the same 1769 consumers

the prevalence of pathological and problem gambling reported in the medical literature.[5] Potenza, Kosten, and Rounsaville (2001) report prevalence rates for pathological gambling and problem gambling of 5.4% and 4.6% respectively amongst casino gamblers. Our results provide some support for these estimates, that are based on survey data.

The coefficient for $LastWin$ at the individual consumer level cannot be distinguished from zero for most of the consumers for both the bet and play sub-models. Thus, at least in this context, we do not find support for the hot hand myth and the gambler's fallacy. This is somewhat contradictory to the evidence for the "Lucky Store" effect found in Guryan and Kearney (2005), where sales for state lottery tickets go up at stores that sold winning tickets in previous periods. One key issue to consider is that in that study, the effect of the winning store on future sales of state lotteries has two competing explanations. The first is that consumers mistakenly believe in the hot hand myth. The other explanation is that a winning ticket generates localized advertising for the store that lottery as a whole and hence sales of the lottery go up in the local region. By contrast, our data are at the individual consumer level and hence our test for the presence of the hot hand myth and the gambler's fallacy is much cleaner.

The coefficient for $CumWin$ is positive and significant for about 30% of the consumers, with a higher degree of significance for the play equation parameters. This provides some weak evidence for the fact that consumers respond to their overall level of wealth gained or lost during the trip. If consumers lose a lot of money, their participation is less likely. The effect on bet amount is not clear *a priori*. We empirically find a positive effect on bet amounts though at the individual consumer level, this parameter is significant only for a minority of consumers.

An interesting corollary of our findings is that they rule out addictive effects for a majority of consumers. We also control for other possible reasons that could explain the propensity to play and bet during a visit. These include previous win/loss, cumulative wins/losses, duration in a spent in a casino and unobserved heterogeneity. Thus, the interesting question is what explains the visit and play behavior of these "non-addicted" consumers. One possible explantion is that the derive consumption utility from the whole casino experience.

[5]See, for instance, Potenza, Kosten, and Rounsaville (2001) for definitions of these terms. Pathological gambling is the more severe form of gambling disorder and is closer to addictive behavior

This is consistent with arguments that have been made by proponents of the industry i.e., that gambling is a just another form of entertainment.

The individual-level coefficient for *Comps* is positive and significant for a large majority of consumers in both the bet and play equations. While the effect of *Comps* is likely to be predicted to be positive on consumer gambling behavior, it was unclear whether they induce consumers to play more or to bet more. Hence, our finding that they influence both the bet amount and the participation decision of consumers provides us important insights about the mechanism through which marketing activities work in this industry.

Table 3 about here

6.3 Elasticities

The parameter estimates discussed above are useful to understand the directional nature of the various effects we have discussed. In order to understand their magnitudes, we compute elasticities for various factors that affect the play and participation decisions and report them in Table 4. These elasticities are computed for the mean consumer and are computed for the covariates for each observation in our data. We then take the average across these observations to find the elasticities that we report in the Table. We report three sets of elasticities for each covariate. The first is the elasticity of the bet amount, conditional on participation. The second is the elasticity of the probability of participation. The third is the total elasticity, which combines the participation and bet amount and is hence the unconditional elasticity of the bet amount.

The unconditional bet amount elasticity for *Comps* is about 0.14. Interestingly, while comps could be expected to work like promotions or coupons, the estimated elasticity suggests that they are more akin to advertising. We base this conclusion on the fact that price and promotions elasticities are typically greater than 1 (in magnitude) while advertising elasticities are typically around 0.10. This is an interesting finding that deserves further investigation. The bet amount elasticity of $LastBet$ is negative and fairly large for the average consumer at -0.21. The major component of this elasticity is from the bet equation. The elasticity of $LastWin$ is insignificant since the parameter estimates are not significant even at the population level. The elasticity of $CumWin$ is relatively small, suggesting that the wealth effect on participation and bet amounts is relatively weak. Finally, $TimeSpent$ has a relatively large negative effect on the participation decision with an elasticity of -0.20.

Table 4 about here

6.4 Parameter correlations

An objective of ours in this study was to see if there are interactions between addiction and marketing effects in this category. In particular, we asked the question whether consumers who are more prone to addiction are also more prone to marketing. Similarly, we wished to see if those who are more prone to the hot hand myth or the gambler's fallacy are also simultaneously more prone to marketing efforts by casinos. This second question is rendered irrelevant by the fact that we do not find evidence for the existence of either effect. We computed the correlations of the individual-level parameters and report these in Tables 5 and 6 for the bet and play sub-models respectively.

Tables 5 and 6 about here

The coefficients for *LastBet* and *Comps* are positively correlated in both the bet and play sub-models, with correlations of 0.1679 and 0.3755 respectively. This provides some weak evidence for the fact that consumers who are more prone to addictive behaviors are also more strongly affected by marketing activities in this category. To illustrate this, we present the scatter plots of the individual-level mean estimates for these two coefficients for the bet and play sub-models respectively in Figures 1 and 2.

We also find a high level of correlation between the coefficient for *LastBet* and the intercept in both the bet and play sub-models. This suggests that consumers who have a high instrinsic bet amount and a high intrinsic probability of participation are also more likely to be prone to addictive behaviors. This can be seen in the scatter plots in figures 3 and 4. Finally, the correlation between the coefficients for *LastBet* and $CumWin$ is negative and large in both sub-models, suggesting that consumers more prone to addiction have a lower wealth effect on their bet amounts as well as participation probabilities. Scatter plots documenting these correlations can be seen in Figures 5 and 6.

Figures 1 - 6 about here

7 Conclusion

Our paper adds to the small but growing body of research that investigates gambling behavior. As noted before, the gambling industry is a large and pervasive part of American society today. This has led to a lot of debate on the benefits and costs of gambling. Our access to a rich and new dataset on individual consumer behavior vis-a-vis casino visitation and activity allows us to take a data based approach to investigating many of the commonly made arguments about these benefits and costs.

We focus our attention on one of the commonly cited costs of gambling i.e., that it leads to addictive behavior (with potentially harmful individual and societal effects). We use the commonly accepted definition of addiction from the economics literature to test for its presence i.e., that current consumption is affected by past consumption. We fit a model of the play decision and bet amount (given play) to data from a consumer panel of casino visitors over a two year period. Our data are at a highly disaggregate level – we look at play decisions within a given trip for individual consumers. Our modeling approach allows us to exploit the rich variation in the data both across and within individuals.

Our results show that, controlling for other reasons that could induce play, only about 7% of all consumers show evidence for addiction (as defined by us). While this proportion may look small, it is consistent with research carried out in other domains that has focussed on casino gamblers. In addition, we find no evidence for either the hot hand myth or the gambler's fallacy. An interesting (though indirect) implication of our analysis is that casino gambling seems to be another form of entertainment for the non-addicted consumers.

Our paper can also be seen as more descriptive in terms of the role of marketing in this industry. We find that marketing activity has a positive effect on the decision to play and the amount to bet. In terms of effect size, comps seem to be more similar to advertising rather than price promotions. Finally, we find some weak evidence that marketing activity is more effective for consumers who exhibit more addictive behavior.

In terms of continuing research, we see a lot of scope for refining our model by allowing institutional details to inform our modeling choices. For example, we understand that marketing dollars in the industry are not delivered randomly to players i.e., comps and offers are targeted strategically. Once we have more detail on the targeting process, we can incorporate that into the model explicitly. In addition, we have not used any information on the various types of marketing offers and their effect on play and bet behavior. We also continue to refine our model primitives as they relate to customer behavior. A major assumption that we have made is that consumers are myopic. While technically challenging, this is an assumption that we could test for. In addition, we have not explicitly tested for threshold effects that have been postulated in previous research. We have also not tested directly whether the gambling industry provides an alternative form of entertainment. It would be valuable if we could incorporate a test for this directly in our model. Finally, we could also enrich our model by looking at variation in customer behavior across and within trips in more detail.

References

Becker, G. S., M. Grossman, and K. M. Murphy (1994): "An Empirical Analysis of Cigarette Addiction," *American Economic Review*, 84(3), 396–418.

Becker, G. S., and K. M. Murphy (1988): "A Theory of Rational Addiction," *Journal of Political Economy*, 96(4), 675–700.

Chaloupka, F. (1991): "Rational Addictive Behavior and Cigarette Smoking," *Journal of Political Economy*, 99(4), 722–742.

Chib, S., and E. Greenberg (1995): "Understanding the Metropolis Hastings Algorithm," *American Statistician*, 49, 327–335.

Clotfelter, C. T., and P. J. Cook (1993): "The 'Gambler's Fallacy' in Lottery Play," *Management Science*, 39(12), 1521–1525.

Gelfand, A. E., and A. F. M. Smith (1990): "Sampling-Based Approaches to Calculating Marginal Densities," *Journal of the American Statistical Association*, 85(410), 398–409.

Gruber, J., and B. Koszegi (2001): "Is Addiction Rational? Theory and Evidence," *Quarterly Journal of Economics*, 116(4), 1261–1303.

Guryan, J., and M. S. Kearney (2005): "Lucky Stores, Gambling and Addiction: Empirical Evidence from State Lottery Sales," Working Paper, University of Chicago.

Heckman, J. J. (1981): "Statistical Models for Discrete Panel Data," in *Structural Analysis of Discrete Data and Econometric Applications*, ed. by C. F. Manski, and D. L. McFadden. MIT Press.

Jones, J. (2004): "Profile of the American Gambler," available at http://www.harrahs.com/about_us/survey/2004_Survey.pdf, last accessed on June 15, 2006.

Kearney, M. S. (2005): "The Economic Winners and Losers of Legalized Gambling," *National Tax Journal*, 58, 281–302.

Olekalns, N., and P. Bardsley (1996): "Rational Addiction to Caffeine: An Analysis of Coffee Consumption," *Journal of Political Economy*, 104(5), 1100–1104.

Pollak, R. A. (1970): "Habit Formation and Dynamic Demand Functions," *Journal of Political Economy*, 78(4), 745–763.

Potenza, M. N., T. R. Kosten, and B. J. Rounsaville (2001): "Pathological Gambling," *Journal of the American Medical Association*, 286(2), 141–144.

Roy, R., P. K. Chintagunta, and S. Haldar (1996): "A Framework for Investigating Habits: "The Hand of the Past," and Heterogeneity in Dynamic Demand Choice," *Marketing Science*, 15(3), 280–299.

Stigler, G. J., and G. S. Becker (1977): "De Gustibus Non Est Disputandum," *American Economic Review*, 67(2), 76–90.

Tanner, T., and W. Wong (1987): "The Calculation of Posterior Distributions by Data Augmentation," *Journal of the American Statistical Association*, 82, 528–549.

A Likelihood Function

For ease of exposition, we use the following abbreviations in this appendix

$$b_{it} = \ln\left(Bet_{it}\right) \tag{A-1}$$

$$P_{it} = Play_{it} \tag{A-2}$$

$$P_{it}^{*} = Play_{it}^{*} \tag{A-3}$$

Also, define the following

$$d_{it} = \begin{pmatrix} b_{it} \\ P_{it}^{*} \end{pmatrix} \tag{A-4}$$

$$g_{it} = \begin{pmatrix} X_{it}\alpha_i \\ Y_{it}\beta_i \end{pmatrix} \tag{A-5}$$

$$\Sigma = \begin{pmatrix} \sigma^2 & \rho\sigma \\ \rho\sigma & 1 \end{pmatrix} \tag{A-6}$$

$$m_\alpha = \operatorname{vec}\left(M_\alpha\right) \tag{A-7}$$

$$m_\beta = \operatorname{vec}\left(M_\beta\right) \tag{A-8}$$

The likelihood function can be written as follows

$$\mathcal{L} \propto \prod_{i=1}^{N} \left[\begin{array}{c} \prod_{t=1}^{T_i} \left[\begin{array}{c} \left[|\Sigma|^{-0.5} \exp\left(-\frac{1}{2}\left(d_{it}-g_{it}\right)' \Sigma^{-1}\left(d_{it}-g_{it}\right)\right)\right]^{P_{it}} \\ \left[\exp\left(-\frac{1}{2}\left(P_{it}^{*}-Y_{it}\beta_i\right)^2\right)\right]^{1-P_{it}} \\ \left[1\left(P_{it}^{*}>0\right)\right]^{P_{it}}\left[1\left(P_{it}^{*}<0\right)\right]^{1-P_{it}} \end{array} \right] \\ |V_\alpha|^{-0.5} \exp\left(-\frac{1}{2}\left(\alpha_i - M_\alpha z_i\right)' V_\alpha^{-1}\left(\alpha_i - M_\alpha z_i\right)\right) \\ |V_\beta|^{-0.5} \exp\left(-\frac{1}{2}\left(\beta_i - M_\beta z_i\right)' V_\beta^{-1}\left(\beta_i - M_\beta z_i\right)\right) \end{array} \right] \tag{A-9}$$

The joint posterior distribution of all parameters is proportional to the product of the likelihood and the prior densities and is given by

$$f\left(\text{parameters}|\text{data}\right) \propto \begin{array}{c} \prod_{i=1}^{N} \left[\begin{array}{c} \prod_{t=1}^{T_i} \left[\begin{array}{c} \left[|\Sigma|^{-0.5} \exp\left(-\frac{1}{2}\left(d_{it}-g_{it}\right)' \Sigma^{-1}\left(d_{it}-g_{it}\right)\right)\right]^{P_{it}} \\ \left[\exp\left(-\frac{1}{2}\left(P_{it}^{*}-Y_{it}\beta_i\right)^2\right)\right]^{1-P_{it}} \\ \left[1\left(P_{it}^{*}>0\right)\right]^{P_{it}}\left[1\left(P_{it}^{*}<0\right)\right]^{1-P_{it}} \end{array} \right] \\ |V_\alpha|^{-0.5} \exp\left(-\frac{1}{2}\left(\alpha_i - M_\alpha z_i\right)' V_\alpha^{-1}\left(\alpha_i - M_\alpha z_i\right)\right) \\ |V_\beta|^{-0.5} \exp\left(-\frac{1}{2}\left(\beta_i - M_\beta z_i\right)' V_\beta^{-1}\left(\beta_i - M_\beta z_i\right)\right) \end{array} \right] \\ |V_\alpha \otimes A^{-1}|^{-0.5} \exp\left(-\frac{1}{2}\left(m_\alpha - \bar{\alpha}\right)' \left(V_\alpha \otimes A^{-1}\right)^{-1}\left(m_\alpha - \bar{\alpha}\right)\right) \\ |V_\beta \otimes B^{-1}|^{-0.5} \exp\left(-\frac{1}{2}\left(m_\beta - \bar{\beta}\right)' \left(V_\beta \otimes B^{-1}\right)^{-1}\left(m_\beta - \bar{\beta}\right)\right) \\ |V_\alpha|^{-\frac{\mu_\alpha + k_1 + 1}{2}} \operatorname{etr}\left(-\frac{1}{2} S_\alpha V_\alpha^{-1}\right) |V_\beta|^{-\frac{\mu_\beta + k_1 + 1}{2}} \operatorname{etr}\left(-\frac{1}{2} S_\beta V_\beta^{-1}\right) \\ \left(\sigma^2\right)^{-\left(\frac{\nu_0}{2}+1\right)} \exp\left(-\frac{\nu_0 s_0^2}{2\sigma^2}\right) \left(\frac{1+\rho}{2}\right)^{\lambda_1 - 1} \left(\frac{1-\rho}{2}\right)^{\lambda_2 - 1} \end{array} \tag{A-10}$$

B Full Conditional Distributions

The full conditional distributions for each parameter vector is obtained by taking out all the terms from the joint posterior distribution in A-10, since the other terms affect only the proportionality constant. We inspect these terms to see if they are from known distribution families. The joint posterior in our case is somewhat atypical because of the selection problem. It may appear that the full conditional distributions of α_i and β_i as well as those for P_{it}^* are not from known distribution families if we just inspect them. However, on closer inspection, it turns out that they can be written as normal distribution.

In order to see this, it is useful to apply a simple trick of converting the joint density of d_{it}[6] into the product of conditional density of $(b_{it}|P_{it}^*)$ and the marginal density of P_{it}^* when writing the full conditional density for α_i. Similarly, we write this joint density as the product of the conditional density of $(P_{it}^*|b_{it})$ and the marginal density of b_{it} when writing the full conditional density for β_i.

For this purpose, it is useful to note that if

$$\begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix} \sim N\left[\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \Gamma = \begin{pmatrix} \Gamma_{11} & \Gamma_{12} \\ \Gamma_{12} & \Gamma_{22} \end{pmatrix}\right] \tag{B-1}$$

then

$$\theta_1|\theta_2 \sim N\left[\mu_1 - V_{11}^{-1}V_{12}\left(\theta_2 - \mu_2\right), V_{11}^{-1}\right] \text{ where } V = \begin{pmatrix} V_{11} & V_{12} \\ V_{12} & V_{22} \end{pmatrix} = \Gamma^{-1} \tag{B-2}$$

Note also that

$$\Sigma^{-1} = \begin{pmatrix} \frac{1}{\sigma^2(1-\rho^2)} & -\frac{\rho}{\sigma(1-\rho^2)} \\ -\frac{\rho}{\sigma(1-\rho^2)} & \frac{1}{(1-\rho^2)} \end{pmatrix} \tag{B-3}$$

Thus

$$b_{it}|P_{it}^* \sim N\left[X_{it}\alpha_i + \sigma\rho\left(P_{it}^* - Y_{it}\beta_i\right), \sigma^2\left(1-\rho^2\right)\right] \tag{B-4}$$

and

$$P_{it}^*|b_{it} \sim N\left[Y_{it}\beta_i + \frac{\rho}{\sigma}\left(b_{it} - X_{it}\alpha_i\right), 1-\rho^2\right] \tag{B-5}$$

Thus, we can write the density of d_{it} (i.e. the joint density of b_{it} and P_{it}^*) in the first line of A-10 in the following two ways

$$f\left(d_{it}|\cdot\right) = f\left(b_{it}|P_{it}^*, \cdot\right) f\left(P_{it}^*|\cdot\right) \tag{B-6}$$

and

$$f\left(d_{it}|\cdot\right) = f\left(P_{it}^*|b_{it}, \cdot\right) f\left(b_{it}|\cdot\right) \tag{B-7}$$

[6]the topmost line in A-10

Then, the full conditional distributions of α_i and β_i start looking like the familiar normal distribution kernels. Thus, we have the following full conditional distributions for the parameters

1. $\alpha_i|\cdot$

   In order to derive the full conditional distribution of α_i, we first rewrite the joint posterior density in A-10 using B-4 and B-6 and then select all the terms that contain α_i. We can ignore the other terms since they affect only the proportionality constant and not the kernel of the density.

$$f(\alpha_i|\cdot) \propto \begin{array}{c} \prod_{t=1}^{T_i} \left[\exp\left(-\frac{1}{2\sigma^2(1-\rho^2)}\left(b_{it} - X_{it}\alpha_i - \sigma\rho\left(P_{it}^* - Y_{it}\beta_i\right)\right)^2\right)\right]^{P_{it}} \\ \exp\left(-\frac{1}{2}\left(\alpha_i - M_\alpha z_i\right)' V_\alpha^{-1}\left(\alpha_i - M_\alpha z_i\right)\right) \end{array} \quad \text{(B-8)}$$

   By rearranging terms in the above expression, we can show that

$$\begin{array}{c} \alpha_i|\cdot \sim N\left[\tilde{\alpha}_i, \left(\tilde{X}_i'\tilde{X}_i + V_\alpha^{-1}\right)^{-1}\right] \\ \tilde{\alpha}_i = \left(\tilde{X}_i'\tilde{X}_i + V_\alpha^{-1}\right)^{-1}\left(\tilde{X}_i'\tilde{b}_i + V_\alpha^{-1}M_\alpha z_i\right) \end{array} \quad \text{(B-9)}$$

   where $\tilde{b}_i$ is the vector obtained by stacking $\frac{\left[b_{it} - \sigma\rho\left(P_{it}^* - Y_{it}\beta_i\right)\right]}{\sqrt{\sigma^2(1-\rho^2)}}$ for all t for which $P_{it} = 1$, and $\tilde{X}_i$ is the matrix obtained by stacking $\frac{X_{it}}{\sqrt{\sigma^2(1-\rho^2)}}$ for all t for which $P_{it} = 1$.

2. $\beta_i|\cdot$

   The derivation of the full conditional distribution of β_i is similar to that for α_i. However, there is one significant difference. In the case of α_i, the full conditional was affected only by those observations where $P_{it} = 1$. In the case of the full conditional distribution of β_i, observations where $P_it = 0$ also enter, but differently from those where $P_{it} = 0$. Rewriting the joint posterior density in A-10 using B-5 and B-7 and selecting the terms that contain β_i gives us the following full conditional density

$$f(\beta_i|\cdot) \propto \prod_{t=1}^{T_i}\left\{\begin{array}{c} \left[\exp\left(-\frac{1}{2(1-\rho^2)}\left(P_{it}^* - Y_{it}\beta_i - \frac{\rho}{\sigma}\left(b_{it} - X_{it}\alpha_i\right)\right)^2\right)\right]^{P_{it}} \\ \left[\exp\left(-\frac{1}{2}\left(P_{it}^* - Y_{it}\beta_i\right)\right)\right]^{1-P_{it}} \end{array}\right\} \quad \text{(B-10)}$$
$$\exp\left(-\frac{1}{2}\left(\beta_i - M_\beta z_i\right)' V_\beta^{-1}\left(\beta_i - M_\beta z_i\right)\right)$$

   The full conditional distibution of β_i can be shown to be

$$\begin{array}{c} \beta_i|\cdot \sim N\left[\tilde{\beta}_i, \left(\tilde{Y}_i'\tilde{Y}_i + \ddot{Y}_i'\ddot{Y}_i + V_\beta^{-1}\right)^{-1}\right] \\ \tilde{\beta}_i = \left(\tilde{Y}_i'\tilde{Y}_i + \ddot{Y}_i'\ddot{Y}_i + V_\beta^{-1}\right)^{-1}\left(\tilde{Y}_i'\tilde{P}_i^* + \ddot{Y}_i'\ddot{P}_i^* + V_\beta^{-1}M_\beta z_i\right) \end{array} \quad \text{(B-11)}$$

where $\tilde{P}_i^*$ is the vector obtained by stacking $\frac{\left[P_{it}^* - \frac{\rho}{\sigma}(b_{it} - X_{it}\alpha_i)\right]}{\sqrt{1-\rho^2}}$ for all t for which $P_{it} = 1$, $\tilde{Y}_i$ is the matrix obtained by stacking $\frac{Y_{it}}{\sqrt{1-\rho^2}}$ for all t for which $P_{it} = 1$, $\ddot{P}_{it}^*$ is the vector of stacked P_{it}^* for all t for which $P_{it} = 0$ and $\ddot{Y}_{it}$ is the matrix obtained by stacking Y_{it} for all t for which $P_{it} = 0$.

3. $m_\alpha = \text{vec}\,(M_\alpha)\,|\cdot$ and $m_\beta = \text{vec}\,(M_\beta)\,|\cdot$

$$\begin{aligned} m_\alpha|\cdot &\sim n\left[\tilde{m}_\alpha, (Z'Z + A)^{-1}\right] \\ \tilde{m}_\alpha = \text{vec}\left(\tilde{M}_\alpha\right) &\text{ where } \tilde{M}_\alpha = (Z'Z + A)^{-1}\left(z'\alpha + A\bar{M}_\alpha\right) \end{aligned} \tag{B-12}$$

Z is the matrix formed by stacking z_i for all i, α is the matrix formed by stacking all α_i and $\bar{M}_\alpha$ is formed by stacking $\bar{\alpha}$ a column at a time.

Similarly,

$$\begin{aligned} m_\beta|\cdot &\sim N\left[\tilde{m}_\beta, (Z'Z + B)^{-1}\right] \\ \tilde{m}_\beta = \text{vec}\left(\tilde{M}_\beta\right) &\text{ where } \tilde{M}_\beta = (Z'Z + B)^{-1}\left(Z'\beta + B\bar{M}_\beta\right) \end{aligned} \tag{B-13}$$

Z is the matrix formed by stacking z_i for all i, β is the matrix formed by stacking all β_i and $\bar{M}_\beta$ is formed by stacking $\bar{\beta}$ a column at a time.

4. $V_\alpha|\cdot$ and $V_\beta|\cdot$

$$\begin{aligned} V_\alpha|\cdot &\sim \text{Inverse Wishart}\left(\mu_\alpha + N, S_\alpha + E_\alpha' E_\alpha\right) \\ E_\alpha &= \alpha - Z\tilde{M}_\alpha + \left(\tilde{M}_\alpha - \bar{M}_\alpha\right)^{-1} A\left(\tilde{M}_\alpha - \bar{M}_\alpha\right) \end{aligned} \tag{B-14}$$

$\tilde{M}_\alpha$ and $\bar{M}_\alpha$ have already been defined in B-12 and N is the total number of consumers.

Similarly,

$$\begin{aligned} V_\beta|\cdot &\sim \text{Inverse Wishart}\left(\mu_\beta + N, S_\beta + E_\beta' E_\beta\right) \\ E_\beta &= \beta - Z\tilde{M}_\beta + \left(\tilde{M}_\beta - \bar{M}_\beta\right)^{-1} B\left(\tilde{M}_\beta - \bar{M}_\beta\right) \end{aligned} \tag{B-15}$$

5. $P *_{it}\,|\cdot$

Using data augmentation (Tanner and Wong 1987), we can treat the vector of P_{it}^* as parameters and make draws for them from their full conditional distribution like in the case of the other parameters Note that when $P_{it} = 1$, P_{it}^* is drawn from a left-truncated normal distribution, while it is drawn from a right-truncated normal distribution when $P_{it} = 0$. Where our specific formulation differs from the standard binomial probit model is that the since we have a bivariate distribution of the errors, there are further restrictions placed on P_{it}^*.

$$P_{it}^*|\cdot \sim \begin{array}{ll} N\left[Y_{it}\beta_i + \frac{\rho}{\sigma}\left(b_{it} - X_{it}\alpha_i\right), 1-\rho^2\right] 1\left(P_{it}^* > 0\right) & \text{when } P_{it} = 1 \\ N\left[Y_{it}\beta_i, 1\right] 1\left(P_{it}^* < 0\right) & \text{when } P_{it} = 0 \end{array} \tag{B-16}$$

6. $\sigma^2|\cdot$ and $\rho|\cdot$

   The full conditional distributions of these two parameters are not from known distribution families and hence we cannot draw from them directly. We use the Random Walk Metropolis Hastings algorithm (Chib and Greenberg 1995) to make draws from these distributions. We use normal candidate densities for this purpose, using standard errors from likelihood estimates of these parameters and a scaling/tuning parameter that was set by maximizing the relative numerical efficiency to fix the variances of these candidate densities. More details are available from the authors on request.

Table 1: Summary statistics of the data

	Median	Mean	SD
Bet Amount ($)	131	448	1367
Previous Win ($)	20	38	319
Total Cum Win ($)	-46	-98	511
Cum Play Time (Min)	72	115	131
Total Play Time (Min)	124	170	165
Comp ($)	0	9	36

Table 2: Population-level Parameters

Parameter	Mean	Std. Dev.	5th Prcntle	Median	95th Prcntle
Bet sub-model:					
Intercept	3.6626	0.0290	3.6100	3.6644	3.7075
LastBet	-0.0496	0.0069	-0.0596	-0.0502	-0.0372
LastWin	-0.0040	0.0156	-0.0307	-0.0031	0.0181
CumWin	0.0268	0.0108	0.0110	0.0252	0.0465
Comps	0.1183	0.0111	0.1022	0.1171	0.1383
Play sub-model:					
Intercept	0.8600	0.0173	0.8318	0.8585	0.8890
LastBet	-0.0136	0.0020	-0.0169	-0.0136	-0.0103
LastWin	-0.0014	0.0050	-0.0091	-0.0017	0.0074
CumWin	0.0126	0.0039	0.0062	0.0126	0.0192
Comps	0.0780	0.0053	0.0696	0.0778	0.0867
TimeSpent	-0.3337	0.0101	-0.3508	-0.3332	-0.3181
Other parameters:					
σ^2	13.7355	0.1862	13.4120	13.7413	14.0278
ρ	0.2467	0.0033	0.2409	0.2468	0.2519

Table 3: Individual-level Parameters

Parameter	Mean	Std. Dev.	5th Prcntle	Median	95th Prcntle
Bet sub-model:					
Intercept	3.6266	0.0741	3.5311	3.6255	3.7440
LastBet	-0.0458	0.0112	-0.0612	-0.0461	-0.0259
LastWin	-0.0112	0.0039	-0.0159	-0.0113	-0.0056
CumWin	0.0281	0.0070	0.0165	0.0282	0.0378
Comps	0.1231	0.0083	0.1138	0.1234	0.1326
Play sub-model:					
Intercept	0.8594	0.0223	0.8302	0.8567	0.9001
LastBet	-0.0138	0.0088	-0.0257	-0.0145	0.0030
LastWin	-0.0010	0.0031	-0.0043	-0.0011	0.0028
CumWin	0.0123	0.0071	-0.0001	0.0130	0.0213
Comps	0.0784	0.0062	0.0725	0.0783	0.0858
TimeSpent	-0.3306	0.0209	-0.3565	-0.3328	-0.2955

Table 4: Elasticities

Elasticities	Bet Amount	Play Probability	Total
LastBet	-0.1858	-0.0292	-0.2149
LastWin	-0.0042	-0.0002	-0.0044
CumWin	0.0275	0.0062	0.0337
Comps	0.1061	0.0351	0.1412
TimeSpent		-0.1984	

Table 5: Correlations between individual level parameters - bet sub-model

	Intercept	LastBet	LastWin	CumWin	Comps
Intercept	1	0.7959	0.8325	-0.8549	-0.1436
LastBet	0.7959	1	0.6715	-0.8908	0.1679
LastWin	0.8325	0.6715	1	-0.7150	-0.0289
CumWin	-0.8549	-0.8908	-0.7150	1	-0.1413
Comps	-0.1436	0.1679	-0.0289	-0.1413	1

Table 6: Correlations between individual level parameters - play sub-model

	Intercept	LastBet	LastWin	CumWin	Comps	TimeSpent
Intercept	1	0.7733	0.2887	-0.7524	0.4737	0.9106
LastBet	0.7733	1	0.1857	-0.6682	0.3775	0.7770
LastWin	0.2887	0.1857	1	-0.1495	0.2186	0.3380
CumWin	-0.7524	-0.6682	-0.1495	1	-0.5390	-0.8230
Comps	0.4737	0.3775	0.2186	-0.5390	1	0.5339
TimeSpent	0.9106	0.7770	0.3380	-0.8230	0.5339	1

Figure 1: Scatterplot of individual level *Comps* and *LastBet* coefficients for bet sub-model

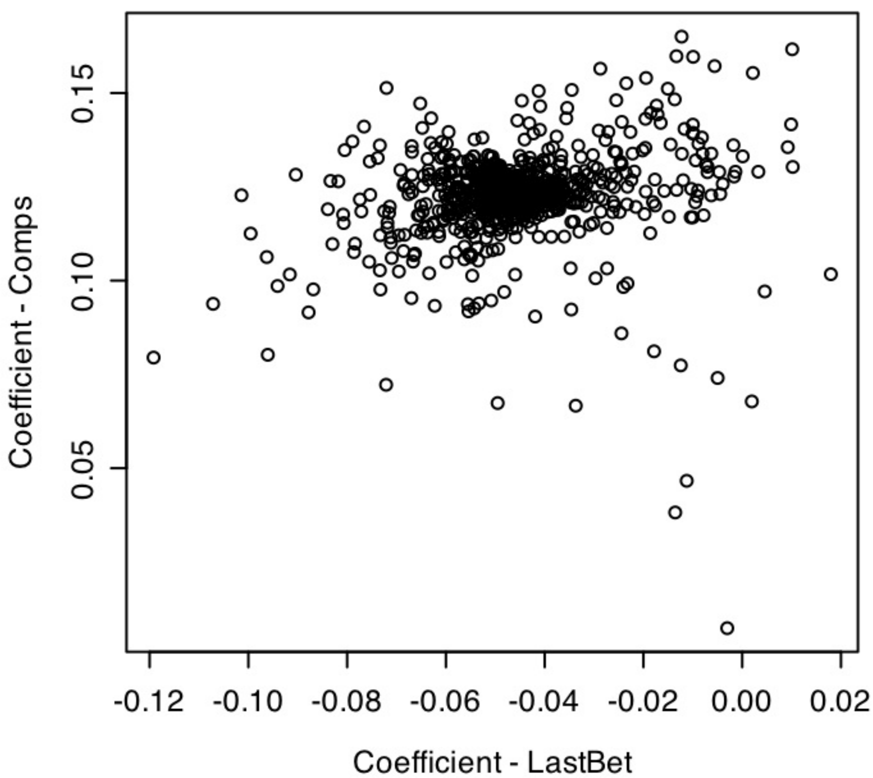

Figure 2: Scatterplot of individual level *Comps* and *LastBet* coefficients for play sub-model

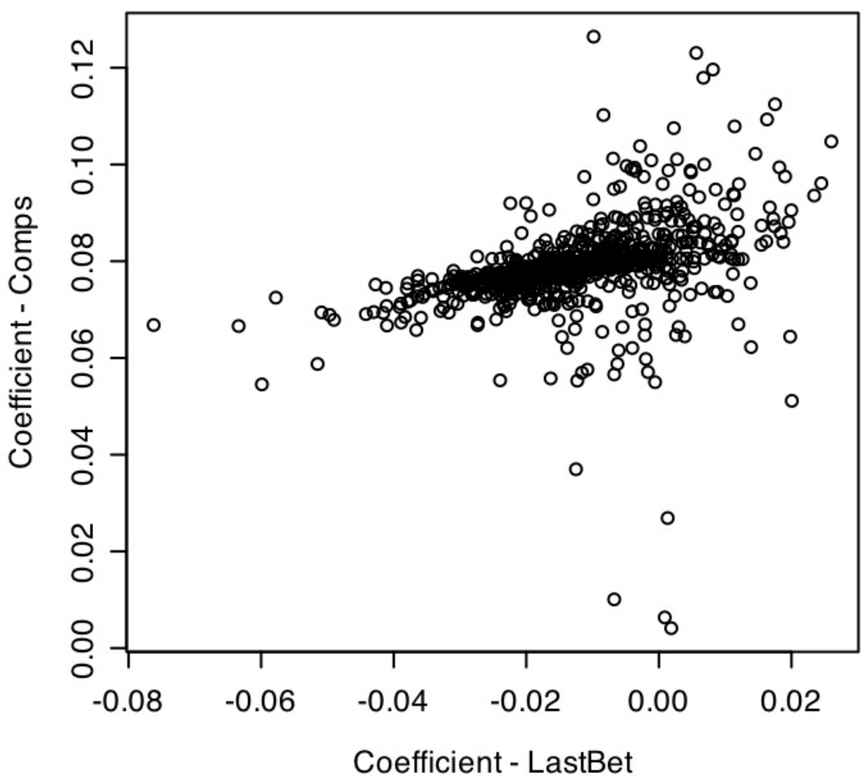

Figure 3: Scatterplot of individual level *Intercepts* and *LastBet* coefficients for bet sub-model

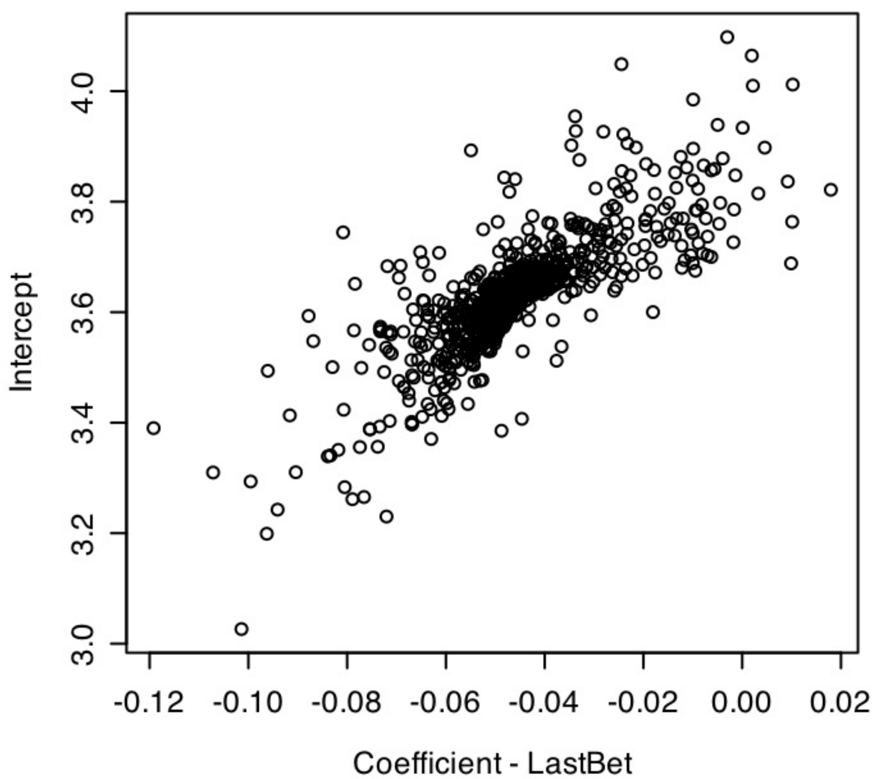

Figure 4: Scatterplot of individual level *Intercepts* and *LastBet* coefficients for play sub-model

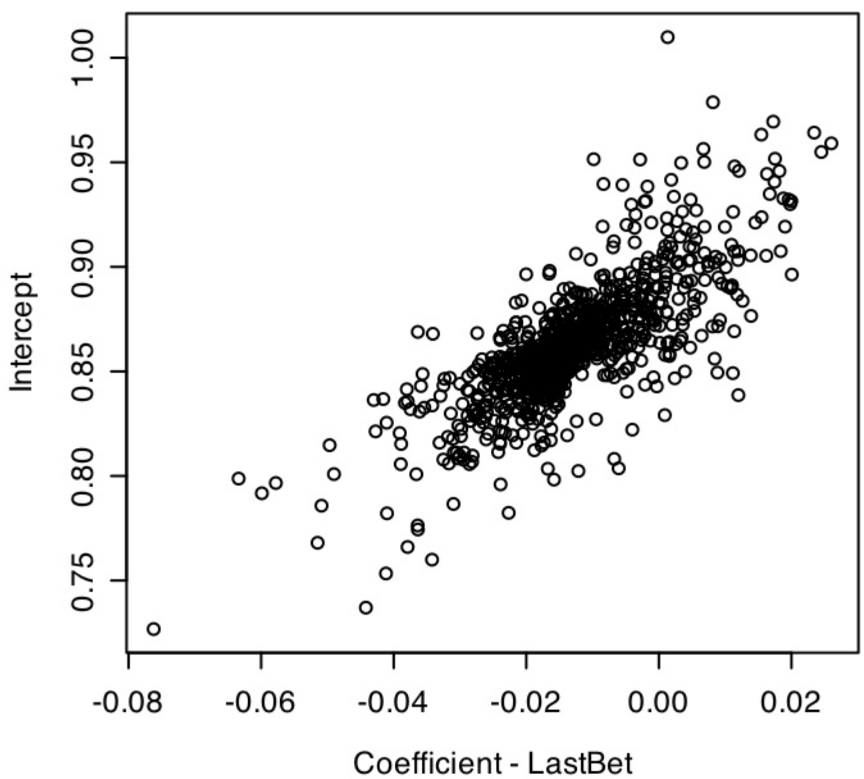

Figure 5: Scatterplot of individual level $CumWin$ and $LastBet$ coefficients for bet sub-model

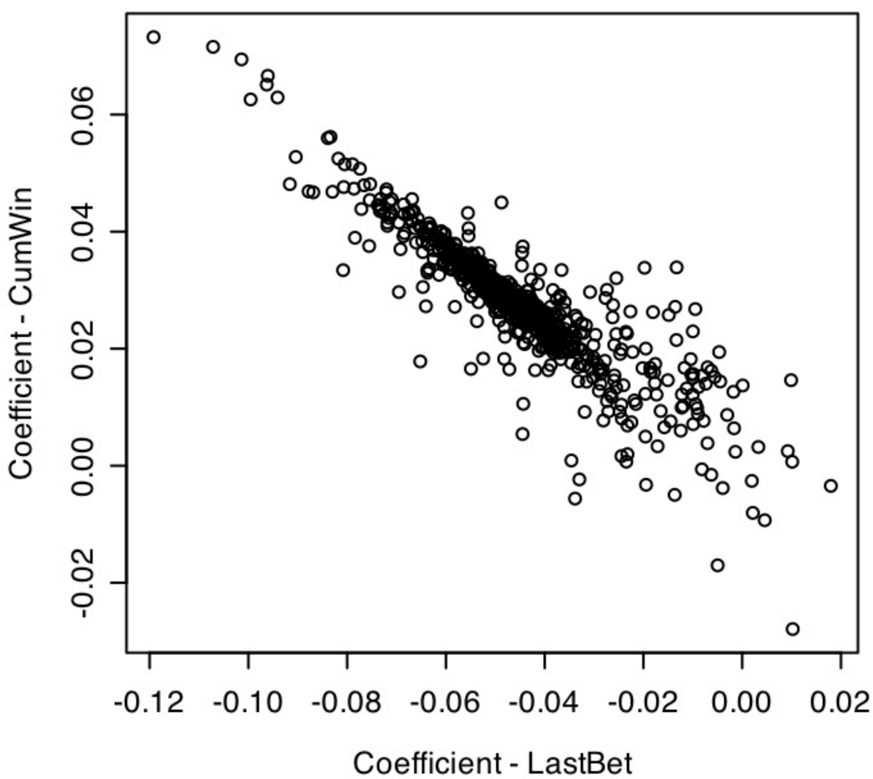

Figure 6: Scatterplot of individual level $CumWin$ and $LastBet$ coefficients for play sub-model