Social Behavior and User Engagement in Competitive Online Gaming: An Empirical Analysis

Liyi Gu¹, Ilya Ryzhov¹, Shawn Mankad², and Bin Han¹

¹University of Maryland
²Cornell University

Abstract

Video games have increasingly taken on a social dimension, with social interaction becoming an important motivating factor for users. Since social gaming helps users derive higher utility from the gaming experience, it may be expected to improve user engagement, retention, and purchase behavior. At the same time, social users may be more vulnerable to churn when one or more individuals from their social network stops playing, and previous work has been mixed on whether, and how, engagement translates to money spent. We conduct an empirical study of the relationship between social interaction and user engagement, retention, and purchase behavior, based on a high-resolution player-level dataset from a major international video game company for one of its premier titles. Our results show that user engagement is highly correlated with certain social dynamics; at the same time, social interaction does not always translate to better retention rates or more purchases. In some cases, high dependence on a small set of friends is positively correlated with churn, indicating a tradeoff between engagement in one title and adoption of others. Early adopters are generally more responsive to the social experience than more casual users.

1 Introduction

Video games represent one of the most popular forms of entertainment among a broad range of individuals. More than 150 million Americans play video games each year, 42% of Americans play at least three hours per week (The Entertainment Software Association, 2015), and 97% of teens in the United States are gamers (Lenhart et al., 2008). In 2014, over 135 million video games were sold to generate over $22 billion in revenue (The Entertainment Software Association, 2015). Two of the largest publicly traded game development companies, Electronic Arts and Activision, have market capitalizations of over $20 billion and annual revenues approaching $5 billion each, as of March 2016.
Social interaction has been widely recognized as an important motivating factor for gamers (Yee, 2006; Cole and Griffiths, 2007). In recent years, the video game industry has undergone a massive transformation that has made the social interaction motive even more important. One major factor is the popularity of gaming on mobile devices (Soh and Tan, 2008); the percentage of gamers that play primarily on their smartphones has grown to 44% in 2014 from 0% of gamers just a decade earlier (The Entertainment Software Association, 2014). Another factor is the emergence of social media platforms, which have themselves become a popular gaming venue: for instance, an estimated 20% of Facebook users play social games (Paavilainen et al., 2013). The “traditional” console (e.g. Microsoft XBox and Sony Playstation) and computer game industry has also been heavily influenced by these developments, and new games on these platforms likewise heavily emphasize social features for communication, institutional formation, competition, and support networks (King et al., 2010). These features are so prominent that cooperative online gameplay often overshadows the single-player features of a game.

The popular press has speculated (Huffington Post, 2014; Bloomberg News, 2013; CNN, 2012) whether console and computer game companies can survive in a market increasingly dominated by smartphones and tablets. In these market conditions of increased competition, a natural question is whether (and how) the social gaming phenomenon can be connected to business objectives such as customer retention and brand management. On the one hand, it seems clear that social interaction is significantly correlated with user engagement and enjoyment. For example, Jansz and Tanis (2007) surveyed gamers and found that the social interaction motive was the “strongest predictor of the time actually spent on gaming.” Lenhart et al. (2008) also found through surveys that 65% of game-playing teens play with other people who are in the room with them, and 27% play games with people who they connect with through the Internet. Playing against human opponents has been shown to positively influence user engagement (Weibel et al., 2008). Additionally, Wohn et al. (2011) found that perceived social interaction was a significant driver of participation in social network games, even if the games themselves lacked features to facilitate such interaction (Consalvo, 2011). Game designers have developed gameplay mechanics for the express purpose of user acquisition and retention; many such mechanics are discussed in Hamari and Järvinen (2011).

On the other hand, it is not clear how well the social interaction motive translates to revenue. Terlutter and Capella (2013) identifies “social factors” as a largely unexplored direction for research on in-game advertising. In the broader context of social networks, Bapna and Umyarov
(2015) has shown that peer influence causes a significant increase in the chances of buying a service. Within social gaming, however, Guo et al. (2015) finds that, while social influence may increase willingness to consume free services, the effect is much weaker when applied to paid content. The empirical literature is mixed on the issue: Cheung et al. (2015) found that “engagement exerted a positive influence on online sales,” whereas Hamari (2015) concluded that enjoyment of a game does make the user more engaged, but actually “reduces the willingness to buy virtual goods.”

In this article, we shed light on these questions with a new empirical analysis of the potential links between social gaming and player engagement, retention, and purchasing behavior for early adopter and casual users. Our analysis is based on player-level data provided by a major international video game company for one of its flagship first-person shooter titles. This game operates through online servers that players connect to using their console or computer to play with other players in real time within a shared game environment. The game has a variety of gameplay modes, though in general the objective is to defeat the other team by scoring the most points, “killing” members of the opposing team, or completing particular objectives, like capturing the opposing team’s base. Players can play multiple rounds on a single gaming server or change servers at any time to play with different players and game settings.

Complete longitudinal data are collected for two cohorts of players, covering every round played by each cohort member (over 882,000 unique game rounds) from product launch in late 2011 through early 2014. Players in the first cohort began playing the game within one month after its release date, whereas players in the second cohort began playing several months after the game’s initial release. For all cohort members, data are available on game usage, game play, and historic behavior that we describe in detail in Section 3. For any round involving at least one cohort member, data are also available for other players that took part; however, their complete play history is unobservable.

From the data, we engineer a set of covariates that characterize different aspects of social behavior by cohort members, and link these social features to user engagement, which we measure using several metrics, mostly related to weekly play time. For each cohort member, we construct a “friends list” containing every player (not necessarily a cohort member) with whom the cohort member played repeatedly. This construction allows us to measure social behavior along multiple dimensions. First, the size of a player’s friends list can be viewed as an absolute measure of social behavior, and thus allows us to test whether more social players
exhibit different dynamics with respect to their playing and purchasing behavior. Second, we
decompose each player’s individual friends list into “casual friends” and “close friends” (where
the precise definition of these terms is player-specific), and subsequently measure the amount of
engagement with each type of friend. This allows us to identify nuanced social effects stemming
from each player’s dependence on casual or close friends for enjoyment of the game.

We find that social behavior is indeed strongly correlated with both the decision to play
each week and the resulting amount of play time. More social players, in both cohorts, exhibit
a higher baseline level of engagement in the game. However, important nuances emerge when
translating engagement to retention and revenue generation. We find that users appear to be
harder to retain when their engagement depends more on playing with close friends, perhaps
because they are more susceptible to churn when those close friends decide to stop playing.
However, such players are also more likely to make more purchases. Conversely, more casual
users, particularly those in the second cohort, make fewer purchases but demonstrate better
long-term engagement in the focus title. Thus, there appears to be a tradeoff between user
engagement in the focus title and user willingness to purchase new titles. We also observe an
intuitive “cannibalization” effect, whereby the release of the next title in the series negatively
impacts engagement and retention for the focus title.

We also find that the two cohorts react differently to the competitive side of the game. The
early adopters in the first cohort demonstrate greater short-term engagement in the focus title
after they perform particularly well, but this does not lead to more purchases, or even to better
long-term retention for the focus title. By contrast, good performance does not seem to spur
the more casual users in the second cohort to become more engaged in the focus title, but does
appear to have a more wide-ranging impact on their interest in gaming and willingness to buy
other titles. This highlights the heterogeneity among gamers’ preferences and may explain some
of the mixed findings in the existing literature.

Thus, while our results verify the link between social interaction and user engagement in
a game, they also imply that game designers should not simply expect highly social users to
buy more. Nor will the encouragement of social behavior necessarily lead to increased revenues.
If the game company’s goal is to encourage early adopters of one title to shift to another, it
may be useful to identify cliques of close friends, who are highly dependent on one another for
engagement, for targeted marketing or special promotions. Alternately, if the goal is to increase
retention of such users in the focus title, it may be useful to encourage them to make new
contacts and social relationships online (Domahidi et al., 2014), to reduce their dependence on close friends. However, if the goal is to strengthen the company’s support among casual players, social behavior appears to be unlikely to offer meaningful guidance. In those cases, it may be more beneficial to focus on the competitive side of the game, perhaps offering promotions or rewards to high-performing players. To support these recommendations, we demonstrate that our models have surprisingly high out-of-sample predictive power, and thus offer a solid basis for managerial action.

We briefly summarize the contributions of our paper. To our knowledge, our work is the first large-scale empirical study to comprehensively examine different dimensions of social behavior in gaming and link them to engagement, retention, and revenue. We find evidence of significant correlations between social behavior and subsequent increases in short-term engagement, which is largely in line with previous work in this area. However, our work is the first to identify significant differences between early adopters and casual users in terms of how this engagement relates to retention and purchases. In particular, we find that increased adoption of other titles comes at the cost of reduced engagement and retention in the focus title. The predictive power of our models also suggests that our proposed measures of social behavior may have significant practical utility, particularly when applied to early adopters.

2 Hypothesis Development and Literature Review

The increasing popularity of video games has generated research interest in a variety of fields, including education (Gee, 2003), economics (Meagher and Teo, 2005), psychology (Sherry et al., 2006), operations (Turner et al., 2011), and information systems (Zhu and Zhang, 2006). Managerial problems studied by this literature include modeling of users’ purchase behavior (Hanner and Zarnekow, 2015), sales forecasting (Yang et al., 2014), utility modeling (Park and Lee, 2011), marketing (Turner et al., 2011), and product innovation (Arakji and Lang, 2007). Much of the existing empirical work is exploratory in nature and based on surveys or interviews; for example, Lehdonvirta (2009) identifies attributes of virtual goods that drive purchase decisions, while Westwood and Griffiths (2010) presents a classification of gamers into six types (one of which is “social gamers”) influenced by different motivational factors.

Our main objective in this study is to use observational, player-specific usage data to understand and improve player engagement, retention and purchasing behavior. The business case for our analysis is supported by the literature on customer engagement, which emphasizes the
influence of customer satisfaction on loyalty, customer retention, and hence profitability (Reichheld, 1992; Voss et al., 2008). Greater customer engagement provides the game company with opportunities to study their more profitable customers, develop better services, and further improve customer loyalty (Mithas et al., 2005). Specific to the game industry, Hsu and Lu (2007) has found that loyalty is positively impacted by enjoyment, while Cheung et al. (2015) presents evidence that increased engagement contributes positively to game sales.

Next, we present four research hypotheses and their theoretical motivations.

**Hypothesis 1:** Social gaming increases playtime and retention rates.

The literature has proposed multiple detailed definitions of customer engagement, which can be measured through surveys or interviews (Cheung et al., 2011, 2015). The game company in our study, however, relies on observational data on players’ in-game behavior. Thus, in our work, we use the player decisions to play and the amount of time spent playing as proxies for engagement. The idea is that play time is positively associated with satisfaction; conversely, a player is likely to have become dissatisfied or lost interest in the game if he or she stops playing it entirely for the remaining time covered by the data.

Under this framework, Hypothesis 1 can be motivated by the large literature on customer satisfaction and loyalty for experience-driven products, where customer engagement can be more important than simply providing quality service (Pullman and Gross, 2004). Deeply engaged customers are more likely to become enthusiastic fans who generate positive word of mouth and increase their spending (Pine and Gilmore, 1999; Vivek et al., 2012). Thus, moving beyond operational factors, like timeliness and dependability (Heim and Sinha, 2001; Mollenkopf et al., 2007; Kumar et al., 2011), user interface navigation (Heim and Sinha, 2001), among others (Srinivasan et al., 2002; Craighead et al., 2004), the goal of the gaming company is to also provide experiences that create “emotionally engaged customers” (Voss et al., 2008).

Studies focusing on the video game industry in particular indicate that social gaming can be the basis for unique and memorable experiences for players (Jansz and Martens, 2005). Peña and Hancock (2006) analyzes text messages from a similar online game and finds that players produce much more socioemotional than task content that is emotional and positive in tone. Sherry et al. (2006) identifies social interaction and peer pressure as the main reasons why many individuals play video games. The potential importance of social gaming to play time and retention relates to key CRM findings that customer satisfaction and retention can be driven from relationship commitments (Gustafsson et al., 2005). In our context, players want
to keep playing as long as their friends are also playing. Moreover, social gaming likely helps
to further differentiate the game from competitors, which can help win loyalty (Seenivasan et al.,
2016). Hsu and Lu (2007) finds that perceived social cohesion, through participation in an
online community, can indirectly influence loyalty to a game company.

Social gaming can also be related to the broad body of work on social influence. For example,
social interactions have been shown to impact participation in the stock-market (Hong et al.,
2004) and in the political process (McClurg, 2003). In the retail setting, Luo (2005) finds
that the mere presence of peers increases the urge for impulsive purchases, and that the effect
increases in magnitude with relationship strength. Siegel (2013) provides several examples from
telecommunications, financial services, and public health illustrating the importance of social
effects in understanding individual behaviors. Kempe et al. (2003) studies how such effects can
be exploited for disseminating ideas or encouraging product adoption; see Bond et al. (2012)
for a case study of political influence through online social networks. We note, however, that
our work also does not belong entirely to this area: competitive online play is task-oriented and
involves a great deal of randomness (e.g., playing against randomly chosen opponents), and does
not lend itself to the kind of social influence analysis performed on, e.g., social networks (Tang
et al., 2009; Ye and Wu, 2010). Rather, we examine the role of social interaction in gaming,
and use the social influence literature to motivate Hypothesis 1.

We also acknowledge some arguments against the hypothesis that social gaming improves
retention rates. For instance, Hu et al. (2016) shows that, when consumers are sensitive to
peer decisions, a product can suddenly become popular due to herding behavior of other social
individuals amplifying adoption decisions. Bapna and Umyarov (2015) provides experimental
evidence that an adopting friend causes a significant increase in the chances of buying a service,
which however may also imply reduced engagement in previous services. Based on these and
other similar findings (Hong et al., 2004; Young, 2009), in our context, if a large portion of
the utility of the gaming experience comes from social gaming, then players may quit the game
simply because enough of their friends have also quit or moved on to a different game.

**Hypothesis 2:** Gamers who experience more social gaming will make more purchases.

It is widely believed that increased customer engagement results in revenue growth, since,
when coupled with data collection, firms have a chance to study their more profitable customers,
design and develop better services, and hence, further improve customer satisfaction and loyalty
(Mithas et al., 2005). Hanner and Zarnekow (2015) finds evidence that users are willing to spend
more money over time as they become more engaged. Similarly, Srinivasan et al. (2002) finds that loyalty in e-commerce has an impact on word-of-mouth promotion, as well as on willingness to pay more. Both of these factors are prominent in social gaming. Cheung et al. (2015) finds a positive link between engagement and sales.

However, counter-arguments are also available. Hamari (2015) finds that enjoyment of the game (a form of engagement) does lead to increased willingness to play the game, but may in fact reduce users’ willingness to buy additional content. Similarly, Guo et al. (2015) finds that, while social influence exerts a significant impact on play time, the corresponding effect on money spent is much weaker. Finally, the weekly time in which a user can play video games may be viewed as a kind of “budget constraint,” reducing any positive effect that engagement may have on spending (Fornell et al., 2010).

**Hypothesis 3:** Social behavior has different effects for early adopters vs. casual gamers. Rogers (2010) defines five categories of adopters of a new service or technology. The first two groups to buy a product, so-called “Innovators” and “Early adopters,” are motivated by social and hedonistic outcomes, that is, these sub-populations make their purchase decision by determining how it will define their social status and also by how much pleasure they can derive from the product (Brown and Venkatesh, 2003). The remaining categories of adopters are analogous to casual gamers in our study, and are driven by other factors, including social influences from their peers and utilitarian judgements about how much the product will improve their overall effectiveness.

An important point is that there are systematic differences in the primary motivations of these different categories; thus, characteristics that appeal to innovators and early adopters may have a different effect on casual users (Moore, 2002). In the context of the gaming industry, Westwood and Griffiths (2010) has also discovered important heterogeneities between “types” (e.g., “social” or “casual”) of gamers, while Drachen et al. (2012) segmented users according to their in-game performance. The existence of these various heterogeneities has important practical implications, e.g., for the design of competitions or game mechanics (Huang et al., 2013), as well as content creation on social media, online forums, official websites, and so on.

**Hypothesis 4:** Playing and purchasing behaviors are sufficiently predictable for managerial action.

In this application, predictive power has important practical consequences. For instance, with sufficient lead time, identifying players who are at risk of stopping their play indefinitely
gives the gaming company opportunities to try targeted marketing campaigns to stop the player from leaving (Siegel, 2013). Our work on this hypothesis relates to the literature on churn that aims to identify customers that are likely to stop using a product or service. Methodologically, we build on past work that applies supervised learning techniques to predict customer churn (Hadden et al., 2007). This literature tends to focus on financial services (Glady et al., 2009), and telecommunications and cable subscriptions (Burez and Van den Poel, 2009). In each of these industries, the firm uses the output of the model to target individuals with promotions that entice them to keep using the product or service. Our work is the first, to our knowledge, to utilize such modeling techniques in the video game context.

3 Data Description

Section 3.1 describes the basic gameplay and user information available in the data. We significantly transformed these data for our study: Section 3.2 describes additional data processing and generation we performed to create attributes describing in-game behavior, and Section 3.3 discusses how we inferred and extracted the social behavior of users from the available information.

3.1 User and gameplay information

The dataset tracks the in-game behavior of 1309 “focal users,” divided into two cohorts of 674 and 635, respectively. Cohort 1 users all began playing the focus game within the first nine weeks of its release, whereas cohort 2 users started after that. Nine weeks is sufficiently far into the life cycle of the product to distinguish between early and late adopters of the game: the first major downloadable expansion pack was released around the time of this cutoff. Both cohorts were monitored for 123 weeks of play, which we refer to as the observation window. The window ends several months after the release of the next game in the series, at which time most users have abandoned the focus game. During the observed time period, five major expansion packs (downloadable content packs or DLCs) were released, each with new game modes and several new levels or “maps.” Throughout the life cycle of the game, users also had the option to purchase a premium account, which allowed greater customization of their in-game characters and provided free access to all future DLCs.

Users play the game by logging on to a server and joining rounds, which are large ongoing battles between two teams. Although it is possible for a team to conclusively “win” or “lose”
a round, this is not the primary incentive of the game, as users may join rounds that are already in progress, leave them before they end, or even switch teams mid-round. Rather, users accumulate points, or *score*, for completing individual in-game tasks, mostly situated around “killing” members of the opposing team. A user is not permanently removed from the round when killed, but rather can “respawn” and continue. In this way, although gameplay is team-based, users are rewarded for their individual performance; a skilled player who happened to be on a losing team is still able to earn a high score and progress in the game.

The dataset contains records of all rounds involving at least one focal user from either cohort. Focal users were selected randomly by the service provider and thus rarely play with each other. Most participants in the rounds are thus non-focal users who were either randomly matched with the focal users, or interacted with the focal users somehow outside the game. Although the dataset includes information about the in-game behavior of non-focal users, this information is limited since we are only able to observe non-focal users when they play with focal users. However, non-focal users are assigned unique IDs, so we are able to identify non-focal users who repeatedly play against the same focal users.

For each participant in a round, we can observe the *time* spent by the user in the round; the user’s in-game *rank* at the start and end of the round, the total *score* earned by the user from the round; and gameplay-specific information such as the game mode and map used for the round, the user’s chosen *role* (offensive, defensive, and in between), any *vehicles* utilized by the user during the round, etc. Multiple metrics of the user’s performance are recorded (such as number of kills and deaths), but some were prone to data collection errors; we focused on combat score as the most reliable metric of performance. The data also included information about teams, but we could not reliably map every participant to a team; furthermore, since players can easily switch teams, we felt that this information was less important for quantifying player skill and performance.

For focal users, we can also see their home country, their signup date on which they first opened their account with the service provider, and the platform they used (PC, PlayStation3 or Xbox 360). Age and gender information is occasionally available, but self reported and unreliable, so we chose not to include it in our models (however, we included random effects to model this and other unobservable information). We also have limited information about *purchases* made by the focal users, most notably whether or not they owned the previous title in the series, and also whether or not they bought a premium account. We also know the number
of products owned by any focal user at the start of the study, which includes all games sold by
the service provider, not just the same series as the focus game. However, subsequent purchases
made during the observation window (with the exception of the premium account for the focus
game) were only tracked in detail for cohort 1. In particular, for cohort 1 users, we can see
whether or not they purchased the next title in the series.

3.2 Data processing and feature generation

We designed and generated a number of attributes to describe user behavior, substantially
transforming the data in the process. Features related to social behavior will be described in
Section 3.3; here, we describe other attributes. First, we aggregated the dataset by week: for
each week and each focal user, we calculated the total play time of the user, i.e., the total
hours spent in rounds during that week. The aggregated dataset is smaller and thus more
computationally tractable; furthermore, aggregation obviates the need to model day-of-week
patterns (e.g., the fact that users generally have more time to play on the weekends).

We also generated additional control covariates describing the user’s performance and in-
game behavior during the week. We used the average score earned per hour of play as a
stand-in for the user’s performance or “skill” level during the week. We also calculated the
proportions of rounds (among all rounds involving the user during the week in question) where
the user employed particular roles or game modes. Similarly we controlled for the proportion
of time spent by the user on maps from various DLCs (some maps are simply more popular
than others), and the proportion of rounds where the user utilized various in-game vehicles. To
further model the effect of new content on player engagement, we defined an indicator variable
for each DLC, which was set to 1 during the four weeks following the DLC’s release date, and
0 in other weeks. Similarly, another indicator variable is set to 1 once the user upgrades to a
premium account. This variable remains at 1 for all weeks after the upgrade was made; if the
player never upgrades, the variable is always zero. For cohort 1, we include similar variables to
describe whether the user owns the next (or the previous) title in the series.

Our statistical models in Section 4 use various measures of user engagement as dependent
variables. These are derived from users’ weekly play time in the following ways: 1) a simple
indicator set to 1 if a particular user plays at all in the given week; 2) an indicator representing
whether or not the user’s weekly play time was unusually high (see Section 4.2 for the exact
definition); 3) the actual play time, in hours, for the week; 4) the last week during the obser-
Figure 1: Empirical distribution of the degree of users’ best friends.

3.3 Features related to social behavior

Because focal users almost never play with each other, and non-focal users are only observable when they play with focal users, the social behavior of focal users must be inferred from repeated contacts with non-focal users. For a given focal user, we constructed a “friends list” containing all the non-focal users who played with that particular focal user on two or more distinct days. We used the number of distinct days, rather than the number of distinct rounds, because the game simply starts a new round when the first round finishes, so multiple mutual rounds within one day may be the result of this mechanism rather than any meaningful social interaction between users.

Throughout this paper, the term “friend” should be interpreted broadly. The users may only be acquainted online (pseudonymously), and may even have met on the game server. Nonetheless, playing the game repeatedly with the same person is a social experience even if the users do not know each other offline. We use “friend” as a catch-all term describing this experience. Many games have built-in social networking features that allow users to explicitly designate other users as friends; such networking data are studied, e.g., by Guo et al. (2015). Unfortunately, the company we partnered with did not keep track of such data for the focus title, and so we chose to infer social behavior directly from the rounds.

Defined in this way, the focal user’s friends can be sorted by degree, which we define to be the number of distinct days when the focal user played with that friend. Degree varies greatly
between friends and also between different friends lists; some focal users have friends whose
degrees are in the hundreds. For each focal user, we define the “best friend” as the friend with
the highest degree; the empirical distribution of this quantity is shown in Figure 1. Comparing
the degree of the best friend across all focal users, we find that 25% of these degrees are between
2 and 4, another 25% are between 5 and 10, another 25% are between 11 and 30, and the final
25% is above 30. We use these quartiles to create “tiers” of social behavior: thus, a focal user
is said to be in tier 0 if the degree of the user’s best friend is 4 or less; tier 1 if the degree of the
user’s best friend is between 5 and 10; and so on.

In this way, tiers measure social behavior in a way that can be compared across users. A
tier-3 user is more social than a tier-0 user, in the sense that the tier-3 user has a richer social
experience involving at least one other person. Although higher-tier users often play more
hours in general, this does not have to be the case: a user who plays twice a month, but with
the same person every time, will belong to a higher tier than a user who plays every day, but
always against random opponents. Similarly, higher-tier users tend to have longer friends lists
in general, but this also does not have to be the case: a user who consistently plays with the
same person may have a higher tier than a user who has a large number of casual friends. We
define the tiers in this way in order to consider social behavior separately from sheer game time:
some players simply have more free time than others and will play the game more overall, but
their level of engagement is not necessarily higher than that of users who only play occasionally,
but derive great enjoyment from the social experience.

Our second measure of social behavior aims to capture such user-specific attributes. For
each focal user, we sort the user’s friends list by degree and divide it in half. The bottom half is
said to be “casual friends,” while the top half is said to be “close friends” (again, these terms are
meant to be interpreted broadly). Then, we calculate the numbers of casual and close friends
with whom the focal user played during each week; in this way, we account for the length of
a user’s friends list as a factor in the user’s social behavior. Furthermore, for each week, we
also calculate the proportions of rounds where the focal user played with casual friends, and a
similar proportion for close friends. Note that the distinction between casual and close friends
is individual for every player: a very social player is likely to have a large friends list and
a high casual/close threshold, whereas a less social player may have a much lower threshold.
Nonetheless, the less social player may still derive significant value from his or her close friends:
consider two hypothetical friends who rarely have time to play, yet derive their enjoyment of
the game primarily from joint sessions (in other words, when they do play, they account for nearly all of each other’s rounds).

In this way, our generated attributes seek to identify users who are more social than others, but also to model instances of social behavior by users who are not highly social compared to others. We also include interaction terms between the indicator variables denoting tiers 1-3, and the proportions of rounds played with casual and close friends, to model the relative impact that playing with close friends has on more vs. less social users. Similarly, we include interactions between the tiers and the numbers of casual and close friends that played with the focal user during the week.

4 Methodology

Section 4.1 presents a model that predicts the simplest possible measure of user engagement, namely whether or not the user played at all in a given time period. We use this model to illustrate the main modeling tools used in our analysis. Then, Section 4.2 discusses other variants that use different metrics of engagement as the response variable.

4.1 Statistical model

Let $I$ be the number of users in a cohort. For week $t = 1, ..., T$, define $y_{i,t} = 1$ if the total play time of user $i$ during week $t$ was strictly greater than zero. Otherwise, let $y_{i,t} = 0$. We formulate a mixed-effect logistic regression model, given by

$$
E(y_{i,t} | b_i) = g^{-1}(x_{i,t-1}^\top \beta + b_i),
$$

where $g$ is the logit link function. The vector $x_{i,t-1}^\top$ contains features that describe the behavior of user $i$ during the previous week. This includes relevant gameplay and purchase information as described previously, as well as our engineered features describing social behavior. We also include user-specific features; for instance, a particular component $x_{i,t-1,k}$ may equal 1 if the user owned a premium account during week $t - 1$. Finally, the time index $t - 1$ is itself a feature in $x_{i,t-1}^\top$, as we expect engagement to decline as the game approaches the end of its life cycle.

The term $b_i$ is a random effect (Laird and Ware, 1982) that models unobservable variation between users. We make the standard assumption that the variables $b_i$ are independently drawn from a common distribution $\mathcal{N}(0, s^2)$, where $s^2$ represents the amount of variation, and
that the individual data points $y_{i,t}$ are conditionally independent given $b_i$. There are several arguments in favour of including random effects in the model. First, $b_i$ models characteristics of user $i$ that influence engagement, but are unknown to the service provider (for example, demographic information). The random effect also controls statistical bias due to multiple data points sharing a common source. Second, random effects explicitly model situations where the given set of users represents a sample of a much larger population, which is certainly the case in our application; estimating the population parameters allows for some degree of inference about the population as a whole. Third, the mixed-effect model has only one additional parameter to estimate (the variance $s^2$), whereas a fixed-effect model would require an additional variable for every user account. The increased size of such a model causes substantial computational difficulties even for moderately-sized datasets. Fourth, by modeling $b_i$ as a random variable, we are expressing the service provider’s inherent uncertainty about individual users and their behavioral patterns. This has the effect of “hedging” our estimates and reducing the risk of overconfidently reporting an effect as being significant in cases where the behavior of the data can be explained by unobservable user characteristics.

Given $\beta$ and $s$, the joint probability of observing $y_{i,t}$ for $t = 1, ..., T$ and $i = 1, ..., I$ is given by

$$L(\beta, s) = \prod_{i=1}^{I} \int_{-\infty}^{\infty} \prod_{t=1}^{T} \left( \frac{e^{x_{i,t-1}^\top \beta + b_i}}{1 + e^{x_{i,t-1}^\top \beta + b_i}} \right)^{y_{i,t}} \left( \frac{1}{1 + e^{x_{i,t-1}^\top \beta + b_i}} \right)^{1-y_{i,t}} \frac{1}{\sqrt{2\pi s^2}} e^{-\frac{b_i^2}{2s^2}} db_i,$$

where the integral represents the expected value of a conditional probability given $b_i$. The maximum-likelihood estimates $(\beta^*, s^*) = \arg \max_{\beta,s} L(\beta, s)$ can be computed using standard statistical software. The statistical significance of the results can then be assessed in the usual ways.

We note that, throughout our analysis, we estimate a separate model for each cohort of users. It is possible to combine the cohorts into a single dataset and expand $\mathbf{x}$ with additional features and interaction terms to model the differences between cohorts. However, this may cause computational difficulties for estimation due to the increased model size. Furthermore, in our case, we are interested in finding out whether the two types of users are influenced by different sets of factors, which can be addressed by estimating two separate models.
4.2 Other measures of engagement

We discuss three additional models that predict different measures of user engagement. On one hand, these models serve as a robustness check for each other and the basic model in Section 4.1; on the other hand, they also allow us to establish more nuanced relationships between engagement and social behavior.

**Binging model.** Let $p_{i,t}$ be the total time that user $i$ spent playing the game during week $t$ (this quantity is also included in the feature vector $\mathbf{x}_{i,t}$ in the basic model). Define $y_{i,t}$ to be 1 if $p_{i,t} \geq q_{i,t}$, where $q_{i,t}$ is the 80th percentile of the set $\{p_{i,t'} : t' = 1, \ldots, t - 1\}$. Otherwise, $y_{i,t} = 0$. We then estimate (1) using this definition of $y_{i,t}$ as the response.

This model predicts whether user $i$ will “binge,” or play for an unusually high length of time, during week $t$. The precise definition of a binge is individualized, since users have different background levels of engagement and some simply play more than others. Furthermore, the definition depends only on the user’s past history up to that week, since our intention is to track whether the user is exhibiting substantially more engagement than was previously typical for him or her. Although the response variable is still related to play time, the proportion of 1s observed is substantially smaller than in the base model from Section 4.1.

**Survival model.** The service provider is particularly interested in understanding when and how users stop playing entirely. We would like to identify any in-game factors that may contribute to dissatisfaction with the service, as well as external factors (such as other purchases, which may “cannibalize” the audience of the present title).

For user $i$, let $T_i = \min \{0 \leq t \leq T : p_{i,t'} = 0 \text{ for all } t \leq t' \leq T\}$ be the last time within the scope of the study that the user was seen to play. Since the overall time horizon $T$ covers most of the life cycle of the game, most users have dropped out by the end. To estimate the distribution of $T_i$ and relate it to user behavior, we formulate a Cox proportional-hazards model (Lin, 2000). This model assumes that $T_i$ is continuous, though in practice it is always estimated from discrete data. The hazard function of $T_i$, given by

$$
\lambda_i(t) = \lim_{dt \to 0} \frac{P(t \leq T_i \leq t + dt \mid T_i \geq t)}{dt},
$$

can be viewed as a measure of how likely user $i$ is to drop out in the next instant in time after $t$. The Cox model assumes

$$
\lambda_i(t) = \Lambda(t) e^{\beta^T \mathbf{x}_i(t) + b_i},
$$

(2)
where $\Lambda$ is a baseline hazard function, $x_i(t)$ is the vector of regression features for user $i$ at time $t$ (note that the features can be time-dependent, as in Fisher and Lin, 1999), and $b_i$ is a user-specific random effect (Sargent, 1998) as before. It is known (Tsiatis and Davidian, 2001) that $\beta$ in (2) can be estimated semiparametrically without the need to specify any explicit form for $\Lambda$; standard statistical packages are available for this purpose (Thomas and Reyes, 2014).

The estimated coefficients have a standard interpretation: if $\beta_k > 0$, then large positive values of $x_{i,k}(t)$ are correlated with a higher chance of dropping out. Unlike our previous logistic regression models, the time index $t$ is not included among the regression features since the dependence of the hazard rate on $t$ is already modeled implicitly using $\Lambda$.

**Tweedie model for weekly play time.** We also consider a model that predicts the quantity $p_{i,t}$, rather than simply the incidence of playing or binging. The Tweedie regression model (Tweedie, 1984) makes the distributional assumption

$$p_{i,t} \sim \sum_{n=1}^{N_{i,t}} \hat{p}_{i,t}^n,$$

where $N_{i,t}$ follows a Poisson distribution, and the random variables $\hat{p}_{i,t}^n$, $n = 1, ..., N_{i,t}$ are independently drawn from a common gamma distribution (by convention, $N_{i,t} = 0$ implies $p_{i,t} = 0$). This distribution is then related to the regression features via the link

$$E(p_{i,t})^{q_1} = x_{i,t-1}^\top \beta, \quad Var(p_{i,t}) = \phi \cdot E(p_{i,t})^{q_2},$$

where $q_1, q_2, \phi$ are additional parameters chosen through maximum-likelihood estimation. We use standard numerical techniques (Dunn and Smyth, 2005) for estimating the regression coefficients $\beta$. To our knowledge, the statistics literature has not developed a mixed-effect version of the Tweedie model.

The Tweedie model is widely used in the insurance industry (De Jong and Heller, 2008) to model random numbers of claims with random amounts. It has several features that are attractive for our application. First, although the response $p_{i,t}$ is continuous, the model allows $P(p_{i,t} = 0) > 0$, making it suitable for zero-inflated data where some users may simply not log on during a given week. Second, (3) has a natural interpretation in the context of weekly play time: $N_{i,t}$ is the number of gaming sessions for user $i$ during week $t$, and the terms $\hat{p}_{i,t}^n$ represent the lengths of individual sessions.
Regression models for purchases. Finally, we consider a set of models that predict purchase behavior. For both cohorts, we have access to the total number of products (not necessarily from the same franchise as the focus game) purchased by each user throughout the duration of the study. We estimate Poisson regression models (McCullagh and Nelder, 1989) that predict this quantity based on the following user characteristics: total play time (during the entire study); total score earned; average growth rate of score; social behavior (tiers 1-3); total number of friends accumulated by the end of the study; the fractions of the user’s total rounds played with casual vs. close friends; whether or not the user bought a premium account for the focus game; whether or not the user owned the previous title in the series; the number of products owned by the user, as well as the total money spent, at the start of the study. In these models, a user represents a single data point, and we do not consider the evolution of the user’s behavior over time, since, in the case of cohort 2, we do not know when the users purchased the products.

5 Results

Section 5.1 presents results for the mixed-effect logistic regression model for predicting a user’s incidence of play in a given week. Section 5.2 presents additional analysis of the other models discussed earlier, covering binging behavior, user churn, and total play time, respectively. Section 5.3 discusses user purchase behavior. Finally, Section 5.4 discusses the predictive power of the models.

5.1 Predicting the incidence of play

We begin with the mixed-effect logistic regression model in (1), where the response $y_{i,t}$ represents the incidence of play by user $i$ at any time during week $t$. Recall that we estimated a separate model for each cohort. Including all of the dummy variables and interaction terms, each model had over 40 features. However, most of the features were not found to be statistically significant for either cohort; due to space considerations, we do not list them here. We also do not report coefficients for features related to the presence of certain DLCs, game modes, maps, or vehicles; while these are important as controls and may be statistically significant, they are less important for our research questions. For instance, we found that the presence of a certain vehicle is negatively correlated with user engagement for cohort 2, but not for cohort 1. Vehicles are quite difficult to control in-game, and it is unsurprising that the more casual users in cohort 2 might not react favourably to them. However, this issue is unrelated to social behavior.
Table 1: Estimated coefficients for logistic regression model predicting the incidence of play.

<table>
<thead>
<tr>
<th>Feature</th>
<th>Cohort 1</th>
<th>p-value</th>
<th>Cohort 2</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>-2.0431</td>
<td>&lt;2e-16</td>
<td>-2.4760</td>
<td>&lt;2e-16</td>
</tr>
<tr>
<td>Index of last week ($t - 1$)</td>
<td>-0.0248</td>
<td>&lt;2e-16</td>
<td>-0.0156</td>
<td>&lt;2e-16</td>
</tr>
<tr>
<td>Hours played last week</td>
<td>0.2818</td>
<td>&lt;2e-16</td>
<td>0.2991</td>
<td>&lt;2e-16</td>
</tr>
<tr>
<td>Skill (score per hour)</td>
<td>0.0014</td>
<td>(not significant)</td>
<td>0.8683</td>
<td></td>
</tr>
<tr>
<td>Premium account</td>
<td>0.4587</td>
<td>&lt;2e-16</td>
<td>0.4832</td>
<td>&lt;2e-16</td>
</tr>
<tr>
<td>% rounds played with casual friends</td>
<td>1.1186</td>
<td>0.0002</td>
<td>1.4550</td>
<td>2.79e-09</td>
</tr>
<tr>
<td>% rounds played with close friends</td>
<td>1.2520</td>
<td>0.0003</td>
<td>1.7910</td>
<td>2.69e-13</td>
</tr>
<tr>
<td>Tier 1 social behavior</td>
<td>0.7066</td>
<td>3.34e-13</td>
<td>0.5706</td>
<td>1.96e-13</td>
</tr>
<tr>
<td>Tier 2 social behavior</td>
<td>0.8694</td>
<td>&lt;2e-16</td>
<td>0.7874</td>
<td>&lt;2e-16</td>
</tr>
<tr>
<td>Tier 3 social behavior</td>
<td>1.5447</td>
<td>&lt;2e-16</td>
<td>1.290</td>
<td>&lt;2e-16</td>
</tr>
<tr>
<td>Tier 1 × % rounds with casual friends</td>
<td>(not significant)</td>
<td>0.3109</td>
<td>-0.6342</td>
<td>0.0256</td>
</tr>
<tr>
<td>Tier 2 × % rounds with casual friends</td>
<td>(not significant)</td>
<td>0.2333</td>
<td>-0.9176</td>
<td>0.0009</td>
</tr>
<tr>
<td>Tier 3 × % rounds with casual friends</td>
<td>-0.7970</td>
<td>0.0101</td>
<td>-0.9112</td>
<td>0.0008</td>
</tr>
<tr>
<td>Tier 2 × % rounds with close friends</td>
<td>(not significant)</td>
<td>0.4732</td>
<td>-0.5920</td>
<td>0.0246</td>
</tr>
<tr>
<td>Tier 3 × % rounds with close friends</td>
<td>(not significant)</td>
<td>0.3627</td>
<td>-0.6019</td>
<td>0.0207</td>
</tr>
<tr>
<td>Owned prior titles in the series</td>
<td>0.1230</td>
<td>0.0003</td>
<td>(unavailable)</td>
<td></td>
</tr>
<tr>
<td>Bought next title in the series</td>
<td>-0.9289</td>
<td>&lt;2e-16</td>
<td>(unavailable)</td>
<td></td>
</tr>
</tbody>
</table>

Table 1 reports results for relevant features that were found to be significant at the 0.05 level in at least one of the cohorts. It can be immediately seen that the results, and even the magnitudes of certain coefficients, are quite similar between cohorts. Some of the results are to be expected: for example, the index $t - 1$ of the previous week has a negative coefficient in both models, implying that engagement decreases as we move closer to the end of the game’s life cycle. The effects of the engineered features are more nuanced and require detailed discussion.

One significant difference is immediately evident: the rate at which the user earned points during week $t - 1$ (recall that this is our stand-in for the user’s current skill level) is strongly significant for cohort 1, but not even approaching significance for cohort 2. This is not a borderline distinction: the p-value is below 0.01 for cohort 1, yet nowhere near 0.05 for cohort 2. As will be shown later, this result is quite consistent throughout all of our models. This answers one of our research questions: early adopters appear to be more engaged by the competitive side of the game, and are more likely to return to it if they perform better. On the other hand, the more casual users in cohort 2 appear to be unaffected (positively or negatively) by their performance.

In marked contrast, social behavior exhibits a strong positive correlation with engagement for both types of users. In both cohorts, the social effect is directionally larger for tier-2 and tier-3 users; however, the magnitude of the effect for a given tier is consistently smaller for more
casual users (cohort 2). Furthermore, both cohorts are always more engaged after playing a greater proportion of their total rounds for the week with friends (both casual and close).

However, different tiers exhibit different sensitivity to the proportion of rounds played with friends, as shown by the interaction terms in Table 1. Moreover, tiers 2 and 3 (the most social users) are the least sensitive to this quantity, although they have higher baseline engagement. Two complementary explanations are possible. First, users who are more social and engaged with the game experience diminishing returns from playing with friends; as they play with more of them, the friends are essentially becoming a commodity. On the other hand, less engaged users with fewer friends overall get more value out of playing with them. Second, our prediction for week $t$ is based on the user’s behavior in week $t - 1$; users who are more social in general should be less sensitive to the precise proportion played with in just the past week. By contrast, less social users may get much more value out of a single positive social experience. Note also that the sensitivity reduction is generally greater for casual friends, as one might expect.

The managerial implication suggested by both explanations is that the most immediate improvement in engagement can be obtained by encouraging less social (lower-tiered) users to play socially more often. The marginal gain of social engagement is higher for these users, and they are at greater risk of dropping out to begin with. Even if the user does not move up in tier, playing more rounds with existing friends improves the odds that the player will continue to be engaged in the game. While higher-tiered users also become more engaged after playing socially, the marginal gain is smaller in comparison.

Finally, the last two rows of Table 1 relate to purchase information, which is only available for cohort 1. Two features of this type were significant. First, owning prior titles in the series (a measure of brand loyalty) unsurprisingly had a positive effect. Second, owning the next title in the series at time $t - 1$ had a negative effect, suggesting that the next title draws committed users away from the current one. For cohort 1, we estimated two different versions of the model, both with and without purchase information; however, for all the features that were common to both models, the estimated coefficients were surprisingly similar (often to the fourth decimal point), and so we do not report both models here.

### 5.2 User engagement: other model variants

We now present results for the models from Section 4.2. As before, for cohort 1 we considered versions of the same model with and without purchase information, but the estimated coefficients
Table 2: Estimated coefficients for logistic regression model predicting the incidence of binging.

<table>
<thead>
<tr>
<th>Feature</th>
<th>Cohort 1 Coefficient</th>
<th>p-value</th>
<th>Cohort 2 Coefficient</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>-6.4281</td>
<td>&lt; 2e-16</td>
<td>-7.7910</td>
<td>&lt; 2e-16</td>
</tr>
<tr>
<td>Index of last week ((t - 1))</td>
<td>-0.0025</td>
<td>0.0089</td>
<td>0.0038</td>
<td>0.0059</td>
</tr>
<tr>
<td>Hours played last week</td>
<td>0.1092</td>
<td>&lt; 2e-16</td>
<td>0.0887</td>
<td>&lt; 2e-16</td>
</tr>
<tr>
<td>Premium account</td>
<td>0.2945</td>
<td>0.0004</td>
<td>(not significant)</td>
<td>0.5897</td>
</tr>
<tr>
<td>No. of casual friends played with</td>
<td>0.0016</td>
<td>0.0032</td>
<td>(not significant)</td>
<td>0.1361</td>
</tr>
<tr>
<td>No. of close friends played with</td>
<td>-0.0016</td>
<td>0.0087</td>
<td>(not significant)</td>
<td>0.1252</td>
</tr>
<tr>
<td>% rounds played with casual friends</td>
<td>2.6790</td>
<td>8.58e-14</td>
<td>2.9770</td>
<td>3.72e-13</td>
</tr>
<tr>
<td>Tier 1 social behavior</td>
<td>0.8784</td>
<td>6.27e-06</td>
<td>0.9319</td>
<td>0.0002</td>
</tr>
<tr>
<td>Tier 2 social behavior</td>
<td>1.0161</td>
<td>1.95e-07</td>
<td>1.3310</td>
<td>4.64e-08</td>
</tr>
<tr>
<td>Tier 3 social behavior</td>
<td>1.2824</td>
<td>1.87e-11</td>
<td>1.4880</td>
<td>3.20e-09</td>
</tr>
<tr>
<td>Tier 1 × % rounds played with casual friends</td>
<td>-0.8399</td>
<td>0.0314</td>
<td>(not significant)</td>
<td>0.1245</td>
</tr>
<tr>
<td>Tier 2 × % rounds played with casual friends</td>
<td>-0.9766</td>
<td>0.0121</td>
<td>-1.1360</td>
<td>0.0142</td>
</tr>
<tr>
<td>Tier 3 × % rounds played with casual friends</td>
<td>-1.5191</td>
<td>6.10e-05</td>
<td>-1.8470</td>
<td>5.29e-05</td>
</tr>
<tr>
<td>Bought next title in the series</td>
<td>-0.9712</td>
<td>7.93e-06</td>
<td>(unavailable)</td>
<td></td>
</tr>
</tbody>
</table>

were so similar for both versions that we only report one set of results. This was the case across all the types of models discussed in Section 4.2.

Binging model. Table 2 reports results for relevant features that were found to be significantly correlated with the incidence of binging. We find that social behavior continues to be strongly correlated with engagement, and the directional effect is greater for tier-2 and tier-3 users compared to less social users.

However, binging appears to be driven by a different type of social behavior. Somewhat counter-intuitively, binging appears to be a casual activity. First, skill level (score per hour), which is related to the competitive aspect of the game, is not identified as being statistically significant for either cohort. Second, interactions with casual friends appear to be much more significant than those with close friends; in fact, playing more with close friends contributes a slight negative effect when it appears at all. As in the previous model, higher-tiered users are less sensitive to the effect of playing with casual friends (since they have a larger number of casual friends overall).

We also find that the presence of a premium account is positively correlated with binging for cohort 1, but not for cohort 2. As will be shown in the next model, the early adopters in cohort 1 are more susceptible to churn, but Table 2 indicates that ownership of a premium account may stimulate more intense short-term engagement in the game. The users in cohort 2 are already more focused on casual enjoyment to begin with, and thus are not significantly affected either way by the premium account.
Table 3: Estimated coefficients for survival model.

<table>
<thead>
<tr>
<th>Feature</th>
<th>Cohort 1</th>
<th>Cohort 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coefficient</td>
<td>p-value</td>
</tr>
<tr>
<td>Hours played last week</td>
<td>0.0376</td>
<td>0.0240</td>
</tr>
<tr>
<td>Tier 1 social behavior</td>
<td>-0.8381</td>
<td>2.9e-05</td>
</tr>
<tr>
<td>Tier 2 social behavior</td>
<td>-1.1082</td>
<td>8.7e-08</td>
</tr>
<tr>
<td>Tier 3 social behavior</td>
<td>-1.5327</td>
<td>2.6e-13</td>
</tr>
<tr>
<td>Tier 1 (\times) % rounds with casual friends</td>
<td>-1.9559</td>
<td>0.0220</td>
</tr>
<tr>
<td>Tier 3 (\times) % rounds with casual friends</td>
<td>-1.8647</td>
<td>0.0200</td>
</tr>
<tr>
<td>Tier 1 (\times) % rounds with close friends</td>
<td>2.1554</td>
<td>0.0230</td>
</tr>
<tr>
<td>Tier 2 (\times) No. of casual friends played with</td>
<td>(not significant)</td>
<td>0.19</td>
</tr>
<tr>
<td>Tier 3 (\times) No. of casual friends played with</td>
<td>(not significant)</td>
<td>0.19</td>
</tr>
<tr>
<td>Tier 1 (\times) No. of close friends played with</td>
<td>-0.0957</td>
<td>0.0480</td>
</tr>
<tr>
<td>Cumulative no. of purchases to date</td>
<td>0.0354</td>
<td>0.0049</td>
</tr>
</tbody>
</table>

Survival model. Table 3 presents results for the Cox survival model from (2). All the versions of this model that we considered had very few statistically significant features. Almost all game modes, vehicles, DLCs and maps appeared to have no significant impact on user churn. User skill and ownership of premium accounts did not appear to help retain users, although they were shown to have a positive impact on short-term engagement in the previous models. On the other hand, social behavior was consistently identified as a significant factor. Note that, in this model, positive coefficients imply a higher likelihood of dropping out, while negative coefficients imply the opposite. Thus, for cohort 1 users, a larger number of cumulative purchases is linked to higher churn for this particular title, indicating a degree of cannibalization between games.

Table 3 suggests several insights. First, users appear to stay in the game longer if they demonstrate more social behavior in general (i.e., if they belong to higher tiers). However, longer retention is linked to more casual behavior: both cohorts experience reduced dropout rates from playing more with casual friends (measured either by the number of such friends or by the proportion of rounds played with them). By contrast, playing more with close friends does not contribute a positive effect anywhere. In fact, for one group of users (tier-1 users in cohort 1), playing with close friends is correlated with higher dropout rates. To explain this behavior, consider a user who spends most of his/her in-game time playing with a small group of close friends. Such a user may be more susceptible to trends; when the user’s friends drop out and move to a different game, the user is more likely to follow them, since the user’s enjoyment of the game is primarily derived from playing it socially. Users may be more vulnerable to such behavior if they are early adopters (cohort 1), and if they are themselves relatively casual (tier-1). However, if the user has more close friends, losing a few of them to a different game
Table 4: Estimated coefficients for Tweedie model.

<table>
<thead>
<tr>
<th>Feature</th>
<th>Cohort 1</th>
<th>Cohort 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>-2.0002</td>
<td>-2.0200</td>
</tr>
<tr>
<td>Index of last week ($t - 1$)</td>
<td>-0.0160</td>
<td>-0.0151</td>
</tr>
<tr>
<td>Hours played last week</td>
<td>0.0846</td>
<td>0.0848</td>
</tr>
<tr>
<td>Skill (score per hour)</td>
<td>0.0010</td>
<td>(not significant)</td>
</tr>
<tr>
<td>Premium account</td>
<td>4.55e-07</td>
<td>0.1910</td>
</tr>
<tr>
<td>No. of close friends played with</td>
<td>0.0413</td>
<td>(not significant)</td>
</tr>
<tr>
<td>% rounds played with casual friends</td>
<td>1.3944</td>
<td>1.9100</td>
</tr>
<tr>
<td>Tier 1 social behavior</td>
<td>0.8707</td>
<td>0.7553</td>
</tr>
<tr>
<td>Tier 2 social behavior</td>
<td>0.9660</td>
<td>0.9244</td>
</tr>
<tr>
<td>Tier 3 social behavior</td>
<td>1.4176</td>
<td>1.3700</td>
</tr>
<tr>
<td>Tier 1 × % rounds with casual friends</td>
<td>(not significant)</td>
<td>-0.7848</td>
</tr>
<tr>
<td>Tier 2 × % rounds with casual friends</td>
<td>0.0920</td>
<td>-1.0590</td>
</tr>
<tr>
<td>Tier 3 × % rounds with casual friends</td>
<td>-1.1086</td>
<td>-0.7726</td>
</tr>
<tr>
<td>Tier 2 × % rounds with close friends</td>
<td>0.0827</td>
<td>-0.9410</td>
</tr>
<tr>
<td>Tier 3 × % rounds with close friends</td>
<td>-1.0845</td>
<td>-0.1020</td>
</tr>
<tr>
<td>Tier 2 × No. of close friends played with</td>
<td>1.3102</td>
<td>4.93e-05</td>
</tr>
<tr>
<td>Tier 3 × No. of close friends played with</td>
<td>0.0396</td>
<td>(not significant)</td>
</tr>
<tr>
<td>Owned prior titles in the series</td>
<td>0.0403</td>
<td>(unavailable)</td>
</tr>
<tr>
<td>Bought next title in the series</td>
<td>-0.6121</td>
<td>(unavailable)</td>
</tr>
<tr>
<td>Number of purchases made</td>
<td>-0.0205</td>
<td>4.13e-10</td>
</tr>
</tbody>
</table>

will not be as important (though there will also be no positive effect from playing with them), moderating the churn effect.

Tweedie model. Recall that the Tweedie model from (3) aims to predict total weekly play time, rather than simply the incidence of play or binging. The results are given in Table 4, and have the standard interpretation that positive coefficients impact play time positively, while the opposite is true for negative coefficients. To our knowledge, no mixed-effect variant of this model is available, but the results can be viewed as a robustness check for the other models in this section.

Overall the results are consistent with those in Table 1: 1) belonging to a higher tier and owning a premium account contributes positively to user engagement; 2) user skill is statistically significant for cohort 1, but not cohort 2; 3) playing more with close friends (as a number or as a proportion of rounds played) is positively correlated with total play time, but much of this effect is canceled out for more social users. Higher-tier users tend to be less sensitive to the effect of playing with both types of friends.

Finally, the purchase information clearly shows the effects of both brand loyalty and canni-
Table 5: Estimated coefficients for Poisson model.

<table>
<thead>
<tr>
<th>Feature</th>
<th>Cohort 1</th>
<th>Cohort 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coefficient</td>
<td>p-value</td>
</tr>
<tr>
<td>(Intercept)</td>
<td>1.2871</td>
<td>&lt; 2e-16</td>
</tr>
<tr>
<td>Total time played</td>
<td>(not significant)</td>
<td>0.4160</td>
</tr>
<tr>
<td>Total score</td>
<td>(not significant)</td>
<td>0.6552</td>
</tr>
<tr>
<td>Avg. skill (score per hour)</td>
<td>(not significant)</td>
<td>0.1041</td>
</tr>
<tr>
<td>Premium account</td>
<td>0.3315</td>
<td>&lt; 2e-16</td>
</tr>
<tr>
<td>Total no. of friends</td>
<td>(not significant)</td>
<td>0.4700</td>
</tr>
<tr>
<td>% rounds played with casual friends</td>
<td>-1.3240</td>
<td>0.0274</td>
</tr>
<tr>
<td>% rounds played with close friends</td>
<td>1.2890</td>
<td>0.0351</td>
</tr>
<tr>
<td>Tier 3 social behavior</td>
<td>-0.8893</td>
<td>3.1e-5</td>
</tr>
<tr>
<td>Tier 2 × % rounds with close friends</td>
<td>1.2921</td>
<td>0.0437</td>
</tr>
<tr>
<td>Tier 3 × % rounds with casual friends</td>
<td>1.2884</td>
<td>0.0417</td>
</tr>
<tr>
<td>Tier 1 × No. of friends</td>
<td>(not significant)</td>
<td>0.5016</td>
</tr>
<tr>
<td>Tier 2 × No. of friends</td>
<td>(not significant)</td>
<td>0.4177</td>
</tr>
<tr>
<td>Tier 3 × No. of friends</td>
<td>(not significant)</td>
<td>0.5007</td>
</tr>
<tr>
<td>No. products owned at start</td>
<td>0.0468</td>
<td>&lt; 2e-16</td>
</tr>
<tr>
<td>Money spent at start</td>
<td>0.0068</td>
<td>2.5e-7</td>
</tr>
<tr>
<td>Bought next title in the series</td>
<td>0.3833</td>
<td>&lt; 2e-16</td>
</tr>
</tbody>
</table>

balization: owning previous titles in the same series implies that the user is more likely to stay engaged with the newer one, but by the same token, buying the next title in the series makes the user more likely to abandon the current one. Owning more games overall also takes away from the user’s engagement in the focus game.

5.3 Discussion of purchase behavior

As discussed in Section 4.2, we also ran Poisson regression models to predict the number of titles purchased by users during the time frame of the study, as well as whether or not they purchased the next title in the series. Recall that, in these models, there is one data point per user, as for cohort 2 we do not have access to detailed information regarding when the titles were purchased. We thus aggregate as many of the user characteristics as possible across weeks; for example, we consider the fraction of the user’s total rounds played with casual vs. close friends (as opposed to last week’s rounds).

Table 5 reports the key results. Unsurprisingly, the volume of purchases made after the beginning of the study is consistently correlated with the volume of prior purchases (likely a reflection of the user’s purchasing power), as well as with the ownership of a premium account or the next title in the series (possibly a brand loyalty effect).

The two cohorts exhibit important differences with regard to social behavior. For Cohort 1
(early adopters), purchases are significantly correlated with various expressions of social behavior. Notably, greater dependence on close friends exhibits a strong positive correlation, whereas a negative correlation is seen for dependence on casual friends. This complements the results of the survival model in Table 3, where the signs were reversed: dependence on close friends was positively linked with the chance of dropping out of the focus game. Thus, Cohort 1 appears to be subject to a tradeoff between engagement in the focus game and adoption of other games; early adopters who play with a group of close friends are susceptible to trends, and can more easily move on to adopt other games.

For Cohort 2 (late adopters), we see almost no correlation between social behavior and purchases. There is a small positive effect for having more friends, but it is largely attenuated at the higher tiers. On the other hand, while the user’s in-game performance (expressed by score and skill level) did not necessarily translate to greater engagement in the focus game (Table 1), it exhibits a strong positive correlation with purchases, suggesting that it may have more wide-ranging impact on the user’s overall interest in gaming.

Overall, these results suggest that social behavior does impact sales, but that this impact is primarily confined to early adopters, and may come at the expense of their engagement in other titles. For late adopters, the best opportunity to increase sales may be to improve their satisfaction with games that they are currently playing, by reinforcing their perception of their performance. In this case the most effective improvements would be related to in-game features rather than social ones.

5.4 Predictive power of models

We now discuss the predictive power of the models we have developed. This is a separate issue from the interpretation of the coefficients, and can help to evaluate the magnitude of the impact of social behavior on player engagement. For the service provider, these results address the question of whether the user population is predictable, as well as the possible utility of the models in identifying engaged users.

As an illustrative example, we consider the logistic regression model from Section 4.1 for predicting the incidence of play. For both cohorts, we use 5-fold cross validation to calculate two widely used performance metrics for predictive models with binary data, namely the AUC, or area under the ROC curve (Smithson and Merkle, 2013), and the lift statistic (Seppänen et al., 2003). The results are shown in Figure 2.
Figures 2(a) and 2(c) suggest that the predictive power of the model is quite high: the AUC, averaged over the five folds of cross-validation, is 0.908 for cohort 1 and 0.893 for cohort 2. The ROC curves exhibit little variation between folds. The binging model from Section 4.2 produces similar results (0.876 for cohort 1, and 0.870 for cohort 2). Moreover, Figures 2(b) and 2(d) suggest that the models would be effective in identifying the users most likely to be engaged in a given week. For example, suppose that service provider wished to identify the 10% of users with the highest chance of logging on sometime during the next week, perhaps in order to offer them promotions. Then, our models would find this group 3 to 4 times more effectively than would a random sample of 10% of the population.
6 Conclusion

We have presented an empirical study of high-resolution player-level data, collected across roughly two years by a major international video game company for one of its flagship first-person shooter titles. From these data, we have engineered features describing various forms of social behavior within that title, and obtained the following insights into the four hypotheses stated in Section 2:

1. User engagement and retention. We find that social behavior is positively correlated with user engagement across both cohorts. However, this type of engagement is short-term in nature and does not always translate to better retention rates. In fact, for early adopters (cohort 1), greater dependence on close friends is linked to greater churn. More casual play is linked to better retention as well as more intense engagement (binging).

2. Purchases. The most consistent predictors of purchases are past purchases. Early adopters do appear to be somewhat more likely to convert their social experience into money spent, but this comes at the cost of their engagement in the focus title. However, identifying such users may be helpful if the company aims to improve engagement in new titles. Later adopters (cohort 2) exhibit a greater link between purchases and past in-game performance.

3. Early vs. late adopters. Building on the previous point, we find significant heterogeneities between early and late adopters, which may help to explain some of the mixed findings in the literature regarding the conversion of engagement to revenue. The two cohorts react differently, not only to the social experience, but also to the competitive side of the game (skill).

4. Predictability. The accuracy displayed in Figure 2 shows that recent in-game behavior has considerable predictive power for future engagement. The company can reliably use these data to identify users for targeted interventions (e.g., discounts or email offers) to improve retention or, conversely, to increase engagement in newer titles.

One surprising managerial implication emerging from our work is that playing with casual friends appears to be better for both short- and long-term engagement. This adds an interesting nuance to the service literature on the importance of creating “emotionally engaged” users who
are enthusiastic fans of the product (Pine and Gilmore, 1999; Vivek et al., 2012). Our results show that in video games there are different types of engagement, and greater variety of play may be more valuable to the company than close-knit social behavior. One potential explanation is that greater casual play creates a wider network with less intense links that is more robust to shocks, since the social network can “absorb” a person leaving. By contrast, users who engage mostly with close friends often have a small but very densely connected social network that is more susceptible to sudden changes. Rather than relying on social behavior, a better strategy for game companies may thus be to pursue efforts to create broader communities and promote involvement at all levels of play.

References


