The Discrepancy Between Video Game Purchase and Consumption in A Content Distribution Network

Yu Ma*
Department of Marketing, Business Econ and Law, University of Alberta, Canada
*yu.ma@ualberta.ca

Keywords: content distribution networks; consumer behavior; video game.

Abstract. This study looked at the user behavior of a content distribution network. We chose a random sample of over 17,000 users from the largest PC video game content distribution network and analyzed their purchase and consumption patterns. We found that an average user purchased 18 games during the 5-year period. However, there was a huge discrepancy between purchase and consumption of these games. An average user would only play half of the purchased games, and the remaining games purchased were never touched. Furthermore our examination showed that the average users (in terms of number of games purchased) had a marginally high utilization rate of their purchases than both light and heavy users. Surprisingly, experience with the service did not increase the utilization rate. These findings have important implications for software developers, publishers and consumer welfare.

Introduction

The video game industry is one of the fastest growing sectors of the entertainment industry. In the US, it reached $22 billion sales in 2014 [1]. As fast and reliable Internet becomes widely available in the world, Internet-based digital content distribution networks (CDN) are rapidly adopted by PC developers and console manufacturers, and are becoming the primary way of software distribution due to its convenience of use, lower cost of distribution than conventional channels, and the ability to limit used game sales and eliminate second hand market. In addition, the popularity of social network encourages all major digital distribution networks to adopt certain functions of social network such as sharing information about user preferences and activities. It leads to the proliferation of behavioral data, publicly available online, and it makes it possible to have a deeper look at game players' behavior (note: the game player is defined as a person who signs up an account with a digital content distribution network, and downloads and plays video game(s). We use the term "player" and "user" interchangeably). A better understanding of the players' behavior can help the software and game developers determine the optimal length of game play, create attractive game designs, and set up ideal releasing schedules for extensions to games such as downloadable content (DLC). Together with traditional sales and profit data, behavior data are used as new source of valuable information for game developers and publishers [2]. It can also help the operators of digital distribution networks to coordinate pricing and promoting activities to increase revenue and profits. Therefore, understanding players' behavior leads to significant values for business intelligence purpose. From the point of consumer welfare, it is also interesting to evaluate consumers' consumption decisions and find out if we can promote healthier video game consumption that can increase the value proposition of video game entertainment to consumers, and ensure the long-term sustainable growth of the entire industry.

Since the proliferation of online behavior data, many studies have harvested data from digital distribution networks or from similar means, and examined user behavior. They primarily focused on how to improve software's ability to recognize patterns in game to create better human like behavior for bots (see [3] for a review). Other studies looked at players' motivation [4] and social groups formation in video games and massive multiplayer online games (MMOs) [5]. However, not much has been done on examining the players' decision of purchase and consumption, i.e., the game playing activities per se. Only a few studies have looked at how individual players played specific

© 2015. The authors - Published by Atlantis Press

games and identified common patterns. An average player’s interest in playing a specific game can be adequately modeled by a Weibull distribution and it is possible to predict when a player stops playing conditioning on having access to the historical data of the player’s behavior in earlier stage of the game and in other games [6]. There was also significant heterogeneity among players. Distinct player groups showed different initial playtime and varying degree of loss of players’ interest in a particular game over time [7]. Our study is one of the first attempts to understand the discrepancy between purchase and consumption. The amount of utilization could depend on many factors, such as whether the purchases match the user’s interest, whether the users have time to play. It is entirely possible for a user to make game purchase decisions based on optimistic estimate of future spare time or imagined gaming experience. After purchase, however, the user may realize that s/he has no time or the playing experience does not align with his/her expectation [8]. Or the user’s preference could change over time due to the goal relevance of the purchase decisions [9].

Data Collection

We harvested publicly available behavior data in July 2015 as described below.

The largest digital distribution platform on PC is developed and maintained by Valve Corporate [10] and called “STEAM”. As of February 2015, STEAM provides access to over 4,500 games, and has over 120 million active users [11]. Although Valve Corporate do not release any official sales figures, it was estimated that STEAM had a 70% share in the PC digital distribution market [12]. STEAM hosts user profiles online for all of its users. If a user is willing to share his/her gaming experience, the user can choose to make the profile public so every visitor to the page can view the user's game inventory, gaming activities, user generated reviews and more. However, there are users who are concerned about privacy. They can set their profile page to private or friends only. Our data collection procedure only includes those users who allow public profile online.

The publicly available profiles provide an opportunity for us to sample from all existing users and examine their game purchase and consumption behavior. Since STEAM assigns a unique 64-bit ID to each user, we first determined the potential range of feasible user IDs, and then randomly generated 160,000 numbers from the range. We then used STEAM application programming interface (API) to retrieve the user information corresponding to the generated numbers [13]. We removed any users who had not been active (playing any game) in the last 6 months to focus on active users only. We also removed any user who did not purchase a game after 2009.

In second step, we iterated over all users in the above list and used STEAM API to collect information about users’ game inventory, which is a list of games owned by the user. In addition, STEAM keeps a record of the total time that a user spent in every owned game since early 2009 [13]. Therefore, we only collected playtime for games released on or after 2010 to ensure STEAM was able to track the complete play history. We also limited our attention in the following analysis only to games released before 2015 to give players enough time to go through their backlogs. Although some users provided basic demographic information in their public profiles such as names, gender, and home locations, majority of the users did not report demographic information, and hence we chose not to rely on demographics for our analysis.

The sample for our empirical analyses contains 17,675 users. The average account age is 3.82 years with a standard deviation of 2.42 years. These users purchased 3,131 unique games released between 2010 and 2014, and spent 5,988,071 hours in total to play these games. This means each player spent roughly 68 hours to play video games per year. Next we look at summary statistics of purchase and consumption and explore the discrepancy.

Purchase and Consumption Discrepancy

Purchases are observed directly from a player's game inventory data. We computed the summary statistics of the game ownership by user: the number of games owned per user account.

On the other hand, consumption is not a well-defined concept since playtime of a game varies significantly across users and across games. For our purpose, we define consumption of a game as a 0/1 discrete state: if a player ever played a game regardless of the actual time spent, the game was consumed. If, by the time of the data collection in July 2015, a game were owned but still not played by a player, we would treat it as if that player would never play it and hence unconsumed. In other words, we define a game as consumed if playtime is non-zero at the time of data collection and as unconsumed if the playtime is zero. Note that we have allowed at least half a year between the release data of the games and the data collection date. It should leave enough time for a player to check out a game if choose to do so.

As a robustness check, we also included a second definition of consumption. Instead of using zero playtime as cutoff for consumed versus unconsumed, we use 1 hour as cutoff. Therefore, in the second definition of consumption, we treat a game as unconsumed if it was played for less than 1 hour in total by a player, and consumed otherwise. Not surprisingly, the correlation between the first and second definition of consumption is very high: $r = .69$ ($p < .001$).

We computed the same summary statistics for the two definitions of game consumptions. Table 1 reports the summary statistics for both the purchase and the consumption.

Table 1 Purchase and Consumption Summary Statistics

	Mean	Std. Dev.	Min	Max
Number of games purchased	18.36	41.96	1	2240
Number of games consumed/ played	12.04	24.87	0	1016
Number of games played for more than 1 hour	8.70	17.93	0	809

On average, a player purchased 18 games in the 5-year period. There was a large variation in the number of games purchased though – a player purchased more than 2,000 games in the 5-year period. Majority of the users purchased fewer than 40 games in the same time period.

The consumption statistics in Table 1, on the other hand, tell a very different story compared to the purchase statistics. First of all, the first definition of consumption outlined earlier is very relaxed and considers a game as consumed if a player spent anytime, even just a minute, in the game. Using this relaxed definition, we found that on average, a player only played 12 out of the 18 purchased games. In other words, only 2/3 of the purchased games were consumed (for at least ran for a minute), and the remaining 1/3 games were never ran or experienced in any way. If we used the second definition of consumption, which is strictly and requires at least one hour of game play to quality a game as consumed, we found that fewer than 9 games out of the 18 purchased games were actually consumed, and half of the purchases were never touched! Further examination showed that the average users had a marginally high utilization rate of their purchases than both light and heavy users. Surprisingly, user experience with the content distribution network did not increase the utilization rate.

Conclusion

In this study we have examined the purchase and consumption discrepancy using a major PC gaming content distribution network. We used a random sample of over 17,000 users from STEAM, a PC gaming distributor. We found that these users spent near 6 million hours playing games from 2010 to 2015. We also found that an average user purchased 18 games during the 5-year period. We discovered a huge discrepancy between purchase and consumption of these games -- an average user would only play half of the purchased games, and the remaining half were never touched.

Since our focal questions are the discrepancy per se, we did not explore the reasons why the discrepancy happened. There are many factors that could contribute to the discrepancy. It could be that users over-estimated the playtime available in the future; it could also be that users did not make the best choices and the purchased games were not really good fit for their preferences. Either way discrepancy could lead to waste and dissatisfaction, and might be harmful for the long-term health of the content distribution network. Future research should examine the underlying cause of the discrepancy and propose strategies to mitigate.

References

[1] ESA (Entertainment Software Association), “Essential facts about the computer and video game industry – 2015 sales, demographc and usage data,” http://www.theesa.com/wp-content/uploads/2015/04/ESA-Essential-Facts-2015.pdf.

[2] A. Drachen and A. Canossa, “Evaluating motion: spatial user behaviour in virtual Environments,” Int. J. Arts and Technology, vol. 4, no. 3, pp. 294–314, 2011.

[3] G. Yannakakis, “Game AI revisited,” in Proc. Conf. Computing Fron- tiers, 2012.

[4] N. Yee, “Motivations for play in online games,” Cyber Psychology & Behavior, vol. 9, no. 6, pp. 772–775, 2006.

[5] C. Thurauand and C.Bauckhage, “Analyzing the evolution of social groups in world of warcraft,” in Proc. IEEE CIG, 2010.

[6] C. Bauckhage, K. Kersting, R. Sifa, C. Thurau, A. Drachen, and A. Canossa, "How players lose interest in playing a game: an empirical study based on distributions of total playing times, " in Proc. IEEE CIG, 2012.

[7] R. Sifa, C. Bauckhaget and A. Drachen, "The playtime principle: large-scale cross-games interest modeling," in Proc. IEEE CIG, 2014.

[8] R. Buehler and D. Griffin. "Planning, personality, and prediction: The role of future focus in optimistic time predictions." Organizational Behavior and Human Decision Processes 92.1 (2003): 80-90

[9] Y. Trope and N. Liberman. "Temporal construal and time-dependent changes in preference." Journal of personality and social psychology 79.6 (2000): 876.

[10] Valve Corporation, http://www.valvesoftware.com/.

[11] R. Smith, "Valve to showcase steamVR hardware, steam machines, & more at GDC 2015", 2015. http://www.anandtech.com/show/9003/valve-to-showcase-steamvr-hardware-steam-machines-more-at-gdc-2015/.

[12] K. Graft, "Stardock reveals impulse, steam market share estimates," 2009. Gamasutra. http://www.gamasutra.com/php-bin/news_index.php?story=26158.

[13] Valve Corporation, “Steam web API,” 2015, https://developer.valvesoftware.com/wiki/Steam_Web_API.