Virtual World Versus Real World: An Economic Study of the Cyber Games Participation

Qingliang Fan[1], Xin Fu[2(✉)], and Shun Cai[2]

[1] Wang Yanan Institute of Studies in Economics, Xiamen University, D312 Economics Building, Xiamen 361005, China
[2] School of Management, Xiamen University, Xiamen 361005, China
xfu@xmu.edu.cn

Abstract. In this paper, we examine the link between the virtual world attributes and real world economic variables. Specifically, we employ three cyber game attributes: playtime, in-game level and achievement (in terms of accomplished 'missions') as the measurements of cyber games participation. We explore the spatial variation of cyber game participation with respect to real world variables including gross regional product per capita, household dispensable income, unemployment rate, number of movie theaters per capita, etc. Moreover, we also study the effects of environmental variables such as precipitation, air pollution etc. on cyber games participation. Using the game data from a very popular cyber game in China and prefecture-level data in 2011, our empirical results show that the income and the availability of other leisure activities are negatively associated with cyber game participation.

Keywords: Cyber games · Consumption behavior · Economic impact · Human capital

1 Introduction

With the growing popularity of cyber games, relatively little attention has been paid to the economic analysis of cyber game participation. This paper aims to study how income level and other economic factors affect the cyber game participation. Different from other goods, such as [9] on coffee consumption and prices in the US, cyber game participation has its unique features. On one hand, playing too much games can deteriorate the player's health and human capital in the long run. On the other hand, cyber games, not different from other leisure activities, could help increase one's productivity. In the seminal paper on the theory of allocation of time, [3] laid the foundation of economic analysis of an economic agent's choice between leisure and production activities.

One of the key features of a virtual world is that the true identify of the players are usually concealed in the network. In this paper, we explore the relationship between cyber game participation and observable economic variables. E.g., how does cyber game participation rate affect local economic development?

F.F.-H. Nah and C.-H. Tan (Eds.): HCIBGO 2017, Part I, LNCS 10293, pp. 58–77, 2017.
DOI: 10.1007/978-3-319-58481-2_6

There might be two different motivations for the game players. First, like other games, cyber game is among the choices of leisure activities such as movies, concerts, etc. This is the relaxing and fun part of the game. Second, they try to 'escape reality'. The latter case raises a severe issue for human capital accumulation and could affect the economic growth in the long run. Cyber games could provide a shelter for some individuals to enjoy the virtual success (as measured by the game performance in our study) as the Chinese labor market is getting increasingly more competitive. We provide empirical evidence that people who have higher income would spend less time in playing cyber games, ceteris paribus. Also, we explore the relationship between cyber game participation with respect to the availability of substitutes such as movie theater, live music concert, performing art, professional sports games, etc., for cities with higher gross regional product per capita (GRPPC) are likely to have more options of after-work social activities. In addition, we explore whether unemployment or bad weather would affect the cyber game participation.

We here discuss briefly the representative features of cyber games that are distinct from other games. Cyber game (also called online games, online PC games, etc.) is a technology which connects players from the world wide web and enables them to play both competitively and cooperatively. Cyber games have the typical features of a virtual world: a synchronous, persistent network of people represented as avatars play different roles and interact with each other in the game. With the advancement of computer technology, cyber games are also becoming increasingly complex (in the perspective of characters design, balance of powers, strategic plays, etc.) and more visually enticing. The popularity of cyber games is accompanied by the rapid development of high speed Internet infrastructure and the increasing number of Internet users. For the consumer base, China has a big population of 'Netizens'. According to a recent report by China Internet Network Information Center, [6], China has approximately 649 million Internet users by 2014, which increases 2.1% than that of 2013, the Internet population rate reaches 47.9%. A big proportion of the Internet user population, especially young generation, play cyber games regularly. Cyber games have a big impact on Chinese youth, especially for male youth, which composed the majority of players population. It is reported in [5] that 65.4% of the players are male, and 66.8% of the players are young people who aged 10–29. They often play cyber games either at home, in Internet café or on smart phones. About 28.5% of these young players have college or higher degrees. Low-income or unwaged players play an important role in the gaming market, it accounts for 27.7% of the player population in 2013.

Gaming industry has developed significantly in China in recent years. The number of players and the market values of gaming industry are growing rapidly. According to the 2014 China Gaming Industry Report (an official report that is released by the Chinese Audio and Digital Game Community (GPC) annually), the number of players in the Chinese gaming market has reached 570 million by 2014, which increases 4.6% than that of 2013. Being the most popular and important game type, online PC game attained its advantages and has received a sales income of 60.89 billion RMB in 2014, which contributes to 53.19% of the total market incomes. Due to the device convenience, players are more likely to play mobile games to fill confetti time (e.g. waiting for a bus and queuing). Conversely, the players of PC games often have higher royalty of the game. The royal players are more willing to spend both time and money to get involved in PC games, purchasing additional virtual equipments and weapons which help players to increase their in-game competency, achieve higher in-game levels, or accomplish more interesting and challenging tasks.

Many researches focus on consumption behavior of conventional goods. In the empirical study of aggregated demand for cigarettes, [4] used a panel data of per capital consumption of cigarettes in state level. [8] studied demand for alcohol using survey data from the National Health Interview Study and find considerable heterogeneity in the price and income elasticities over the full range of the conditional distribution. [2] studied the rational consumption of tobacco using GMM method. [1] empirically studied the consumption of milk. Our study is the first to study how real world economic variables affect cyber game attributes using a novel dataset.

There would be potential endogenous problems with the income variable, since players of low income level would play more cyber games which in turn would affect future income. We employ the instrumental variable (IV) model to address this issue. We choose the freight volume per capita as the instrumental variable for income. Freight volume reflects the economic development of the region, we find a very strong positive correlation between the freight volume and average income of the region in our sample. On the other hand, the freight volume can only affect the behavior of playing cyber games indirectly and through the channel of income. Besides, the gaming data we have is gameplay telemetry in a month, while most of the real world economic variables are reported monthly and even annually. By taking the average of the daily cyber game data, we assume implicitly that the current level of game attributes are comparable among different prefectures. We also use quantile model to study the impact of control variable on different quantiles of cyber game participation.

The paper is organized as follows. In Sect. 2 we introduce the theoretical rational addiction model of the cyber game and the empirical model. Section 3 describes in details the data sources and variables. In Sect. 4 we provide and discuss the empirical results. Section 5 concludes this work and points out future work directions.

2 Model and Preliminaries

2.1 A Basic Economic Model

We here consider a very basic economic model. Assume the utility of an individual depends on two types of goods, the leisure activity goods A, here specifically the cyber games, and the non-leisure goods N. For simplicity, we assume the consumption set is composed of two goods, A and N, and N can be think of as a basket of goods other than A. We also assume the homogeneity of cyber games even though each game is not a perfect substitute of another game and continuity of the excess demand function. Moreover, the current utility also depends on the past consumption level of A, but not for N. The rational consumer's problem is to maximize the sum of lifetime utility discounted at a constant rate r, and denote $\beta = \frac{1}{1+r}$:

$$\sum_{t=1}^{\infty} \beta^{t-1} u(A_t, N_t, S_t) \tag{1}$$

where S is the stock of consumption capital of the individual, $u(\cdot)$ is a strongly concave function of A, N and S, and the past consumption of A affects current utility through the channel of 'learning by doing':

$$\Delta S = A_t - \delta S_t - h(D_t) \tag{2}$$

where $\Delta S = S_{t+1} - S_t$, δ is the instantaneous depreciation rate which measures the exogenous rate of disappearance of the physical and mental effects of past consumption of A, and D_t represents expenditures on endogenous depreciation or appreciation and $h(\cdot)$ is a real valued function. $C_t = N_t + P_t A_t$. Suppose I_0 is the initial value of wealth, C_0 is the initial condition indicating the level of goods A's consumption level at time 0. The utility maximization problem is subject to the following constraints

$$\sum_{t=1}^{\infty} \beta^{t-1} (N_t + P_t A_t) = I_0 \tag{3}$$

The first order conditions satisfy that the marginal utility of wealth is equal to the marginal utility of consumption of N_t in each period, denote $u(t) = u(A_t, N_t, S_t)$:

$$\frac{\partial u(t)}{\partial N_t} = I_t \tag{4}$$

And the marginal utility of current consumption of A, u_1 plus the next period's utility of today's consumption, u_2, equals $I_t P_t$, which is the marginal utility of wealth multiplies the current price of A:

$$\frac{\partial u(t)}{\partial A_t} + \beta \frac{\partial u(t+1)}{\partial A_t} = I_t \tag{5}$$

Assume that for a heterogeneous population H in each area (prefecture, in our case, and for simplicity of the notations we ignore the subscript i) there exists a finite partition $\{\chi_k\}_{k \in K}$ of the set $\mathbb{R}_+^H$ of all conceivable income assignments and for every $k \in K$ there is an aggregate demand function $F^k(P, G)$, where G denotes an income distribution function, such that

$$\frac{1}{|H|} \sum_{h \in H} f^h(P, x^h) = F^k(P, G_{x^h}) \tag{6}$$

for every nonnegative income assignment $x^h_{h \in H}$ in the set χ_k and for every $P \in \mathbb{P}$ which is any strictly positive price vector, and where $|H|$ is the number of population, $f^h(P, x^h)$ is the demand function for each individual $h \in H$ which is derived from the individual utility maximization problem.

2.2 The Empirical Model

The basic empirical model we consider in this section is

$$y_i = \alpha_i + \beta inc_i + \gamma X_i + \epsilon_i \tag{7}$$

where y_i is the variable of cyber game participation. We use three different variables to study the participation of cyber games in different dimensions. The playing time which is considered as the proxy for cyber game participation, the in-game level variable and achievement which is measured by the number of missions accomplished are the proxy for participation and performance of playing, respectively. The inc_i is dispensable income level of prefecture i, we also use GRPPC as income level proxy, X_i is a vector of other exogenous variables which are considered as the determinant of the cyber game variables, these variables include substitute goods, Internet penetration rate, unemployment rate, environmental variables such as rainfall, air pollution, etc.

Since the income variable could potentially affect other unobserved factors that also affect the average cyber games variable, we use the freight volume as an instrumental variable for income. From the economic model of cyber game participation, the income and other control variables would have different effects on different quantile of participation level. For example, for the higher quantile of playtime regions, there would be different preferences or habits and that would result in different impact of income. It is very plausible that the different quantile of playtime would response differently to the economic welfare and environmental variables. We therefore estimate a quantile regression model to identify the heterogeneous effects of income and other control variables across regions. The variables we consider in the quantile regression model is the same as in the aforementioned linear model [7].

3 Data Description

3.1 Dragon Nest

We use a realistic cyber game dataset from Dragon Nest[1] to investigate the proposed model. Dragon Nest is a free-to-play action Massive Multiplayer Online Role Playing Game (MMORPG) developed by *Eyedentity Games*[2]. The game incorporates a non-targeting combat system which provides players a fast paced action filled experience. Players have complete control over every single movement of their chosen character. The players are allowed to choose from a range of playable characters (e.g. Warrior, Archer, Cleric and Academic, etc.) to defeat

1 http://dn.sdo.com/web9/home/index.asp.
2 http://www.eyedentitygames.com/main.asp.

diverse monsters and other players. One player can have one or many characters in the game. Dragon Nest allows the players to initially play for free, but the players are encouraged to speed money in the game to explore and enjoy more game features, such as customized gear for the characters, mainly to increase the visual attractiveness for the characters.

Throughout the game, a number of designed tasks are provided in dungeons. In most MMORPGs, dungeon (often interchanges with the term *instance*) is defined as a private area where allows characters to team up and complete tasks without interference from other characters[3]. In dungeons, characters fight against groups of monsters using their skills and weapons. Dragon Nest also provides some more difficult and challenging tasks (known as missions in this game) for advanced players. Once the player accomplishes the mission, special honors or achievements will be given. Such honers/achievements would help to differentiate the player from other regular players and would be beneficial to enhance the player's reputation and influence in the virtual world. Accomplishing missions not only would help players to gain playing skills, but also to upgrade their in-game levels. Note that, Dragon Nest has an independent skill system, which indicates that a high level player is not necessarily a skillful player. A more experienced player that with a lower level can also overcome a less experienced player but with a higher level using his skills.

The Chinese version of Dragon Nest is published and operated by *Shengda Games*[4], one of the biggest and most successful game operators in China, since July 2010. In a few years, Dragon Nest has become one of the leading games of Shengda. Within 3 days of the official release, the number of Peak Concurrent Users (PCU) reaches 700,000[5]. In gaming industry, to attract and retain more players, game operators often employ some marketing strategies for product advertisements and promotions. The common promotion strategies include advertise the product in popular and authorized online game media (e.g. 17173[6] is renowned as the top game website in China), build player forum in the product official website and organize competitions regularly. Besides these online promotions, some offline marketing promotions are often employed by operators, especially at the beginning of the product launch. Practically, such regional promotions are often started from where the game operators geographically based. In the case of Dragon Nest, since *Shengda Games* is based in Shanghai, several promotions were held in Shanghai and the nearby cities.

3.2 The Cyber Game Data Set

The collected dataset contain one-month gameplay telemetry, from 1^{st} March 2011 to 31^{st} March 2011. In total, the game participation (including login details, playtime, social activities, achieved levels, coins, etc.) of $4,811,925$ unique characters are collected. To reveal the relationship between the virtual world success

[3] http://www.wowwiki.com.
[4] http://www.shandagames.com/us-en/index.html.
[5] http://dn.178.com/201101/89835568248.html.
[6] http://www.17173.com/.

and real-world economic well beings, this study extracts the related information from two perspectives: (1) data from the virtual world - attributes that are associated with players' engagement and their in-game performances; (2) data from the real-world - for the given dataset, the unique attribute that relates to the real-world is the log-in IP. In this work, the virtual world success is measured by three representative gameplay attributes, namely the playtime, in-game level and achievement.

The raw dataset records every single login details, including character ID, login IP address, login timestamp and logout timestamp. Using this game log, the in-game playtime for a given character can be derived. As aforementioned, by accomplishing missions, players can upgrade their in-game level, the game server will automatically capture the every piece of game level update information (e.g. character ID, current level, new level, upgrade timestamp and map ID). The collected dataset includes the information of the completed missions, and the obtained achievements are also recorded. The three chosen gameplay attributes are defined as:

- playtime: this attribute refers to the aggregated game playtime for a given character during the observed time period (i.e. March 2011). Given a character, this attribute is derived from the login log by calculating each login time length and then aggregating them. The derived result is represented in Hour unit.
- in-game level: this attribute represents the achieved in-game level (i.e. the maximum recorded level) of a given character by the end of the observed time period (i.e. 31^{st} March 2011). Note that, this attribute just indicates the current in-game level, but does not mean that the in-game levels are all achieved within the observed time period, as the level may be carried from previous months. In the released version of March 2011, the maximum in-game level of is 40. However, since Dragon Nest is a dungeon-driven game, reaching the maximum in-game level is not the end of the game. The main purpose of the game level-up is more like a tutorial to train the players using different skills and weapons, help them to get familiar with the game story and settings, and provide relative simple tasks for them to gain experiences and increase competency. In so doing, some players can quickly reach the maximum level, and then apply the obtained experiences and skills to accomplish more challenging missions in different dungeons.
- achievement: this attribute refers to the achievements that a given character obtained during the observed time period. The achievement is represented by the sum of the missions that a character accomplished in March 2011. It is important to note that the missions in Dragon Nest are with different difficulty levels. Obviously, the difficult mission would consume more time and resources to complete. However, due to the availability of the collected dataset, there is no associated information to distinguish the easy missions and difficult missions. In this work, the difficulty levels of completed missions are treated equally, only the number of accomplished missions are taken into account. This is a weakness of this work, and further investigations are required when the information of difficulty level become available.

Besides the attributes from the virtual world, we also use the Chinese IP library to convert the raw IP addresses captured in the login log to geo locations. This enables us to build the linkage between a virtual character and the location of real-life player, and the identified geo locations provide the basis for the further analysis of the related economic data.

3.3 Cyber Game Data Preprocessing

The raw data set contains $4,811,925$ unique characters and $11,552,998$ login records. Compared to other application domains, MMORPGs often result in relatively low player retention rate. For the vast majority of games, a significant proportion of players who signed up but never play the game at all. It has been reported that approximately 85% of players do not return after the first day, and game company should expect to lose 96% of their user base within 12 months [7]. This indicates that the raw data set includes loads of 'never-play' gamers information, but such information is beyond our research interests. For data preprocessing, first, this work employs a two-level filtering mechanism to filter out these players and select the ones who have spent sufficient amount of time to play the game. After the following filtering step, the number of characters reduces from $4,811,925$ to $64,448$. Specifically, we process the data as following:

- login level filtering: only the login duration that is longer than 10 min is regarded as a valid game-play login. This is because MMOPGs often require heavy user involvements than other types of games (e.g. mobile games). A simple task at least requires 5–10 min to complete and a dungeon would require longer. The raw login log may also capture some non-sense login details (e.g. the player suddenly loses the Internet connection, and the login duration is only a few seconds), and this step greatly reduces the volume of login log and would speed up the further computations and analysis.
- character level filtering: the login information will be aggregated according to the character IDs. Within the observed time period, the characters who either reach 60 login times or 30 h accumulative login time length will be retained for further analysis.

Second, for a given character, in principle, he/she may use different IP address to login the game. This may cause that the behavior of a single character being multiple counted when analyzing the collective behaviors for a given geo location. To address this, for each character, we initially list all identified geo locations that he/she has used for login, and then count the number of their login times. In this work, it is assumed that the most frequently used geo location is the place where a player actually based in the real-world. Hence, a one-to-one mapping table is derived to store the character ID and his/her the most frequent login geo location. In so doing, a character only contributes to the collective user behavior for a single geo location. In our data set we also have players from foreign countries outside China. Since the main focus of this paper is to study the relationship between playing game behaviors and Chinese prefecture variables, we deleted those players whose IP address is not from China. To reveal the relationship between virtual world success and real world economic well beings, all the in-game behaviors will be accumulated according to different geo locations. Given a geo location, the total number of based characters and their aggregated in-game performance will be calculated. Such data will then be used to derive the average playtime, in-game levels and achievement for a geo location. A summary of the pre-processed attributes is shown in Table 1. Figures 1 and 2 plot the geographic distributions of average play time and regional economic development in China, March, 2011.

Table 1. Summary of attributes

Dimensions	Attributes	Description
Data from the virtual world	Playtime	The accumulative game playing time for a given character
	Level	Achieved in-game level for a given character (min = 0, max = 40)
	Achievement	# of mission accomplished for a given character (min = 1, max = 675)
Data from the real-world	IP address	The originally recorded login IP address
	Geo location	The converted geo location for a given IP address
Aggregated data	Sum_character	# of included distinct characters in a given geo location
	Average_playtime	Given a geo location, the average playtime, it is derived from total_playtime/Sum_character (min = 0.17 h, max = 235.81 h)
	Average_level	Given a geo location, the average achieved in-game level, it is derived from total_level/Sum_character (min = 3, max = 40)
	Average_achievement	Given a geo location, the average number of accomplished missions, it is derived from total_mission/Sum_character (min = 1, max = 328)

At last, we sort the data according to prefectures and take the average of attributes as total quantity divided by the number of players in the prefecture.

3.4 Economic Data

Economic variables are from China Statistical Yearbook for Regional Economy (2012) and China City Statistical Yearbook (2012). The National Bureau of Statistics of China (NBSC) collects data from each prefecture and publishes them altogether each year. These two yearbooks have collected major social and economic indicators of China in 2011. We use GRPPC (gross regional product per capita) and per capita disposable income of urban households (Income) as proxy for income. The two variables are measured in the unit of 1000 RMB. We also control unemployment rate, Internet penetration rate, and the number of cinema per hundred people in the studied area. We use freight volume per capita as instrumental variable for income.

Meteorological data are from China Meteorological Data Sharing Service System of China Meteorological Administration. Data are reported by meteorological stations. We choose the data of each station to represent the level of its located prefecture. Data are collected on monthly precipitation and the number of days when daily precipitation is larger than 0.1 mm in March, 2011 (Days $>$ 0.1 mm). Environmental variables are from China City Statistical Yearbook (2012). Data are collected on average API (air pollution index) which is measured as the arithmetic average of daily reported API of each city in March, 2011.

3.5 Summary statistics

Table 2 contains means, standard deviations and other statistics of the primary variables in the data set. The definitions of variables are given in the previous subsection.

Table 2. Summary Statistics

	Mean	Std. dev.	Median	Min	Max	Sample size
Login time	26.38	7.88	25.37	14.05	51.24	108
Mission	78.15	8.10	77.96	61.00	98.00	113
Level	26.41	2.91	26.12	20.68	32.25	112
Income	22.54	5.52	20.47	14.10	36.51	113
GRPPC	51.73	25.83	45.48	16.39	133.30	113
Cinema	3.77	3.31	2.75	0.43	17.65	113
Internet	37.21	17.12	34.88	10.18	74.95	113
Unemployment	3.18	0.81	3.30	0.70	5.40	113
Precipitation	26.21	30.04	15.90	0.00	135.50	113
Days $>$ 0.1 mm	6.35	5.63	5	0	23	113
API	69.15	11.97	69.00	26.61	117.00	113

Notes: Income and GRPPC are measured in 1000 RMB, Internet and Unemployment are in %, and API is measured in $\mu g/m^3$

4 Empirical Results

In Table 3, we report the OLS results of login time on income and other control variables. We find that income level is negatively associated with cyber game participation. As the average income level increases by 1%, the average play time reduces by about 0.34 h (about 20 min) as shown in column (1). This result is robust to both income proxy variables, dispensable income and GRPPC. It shows that the unbalanced regional economic development level has significant impact on the cyber games participation. For high income regions, the opportunity cost

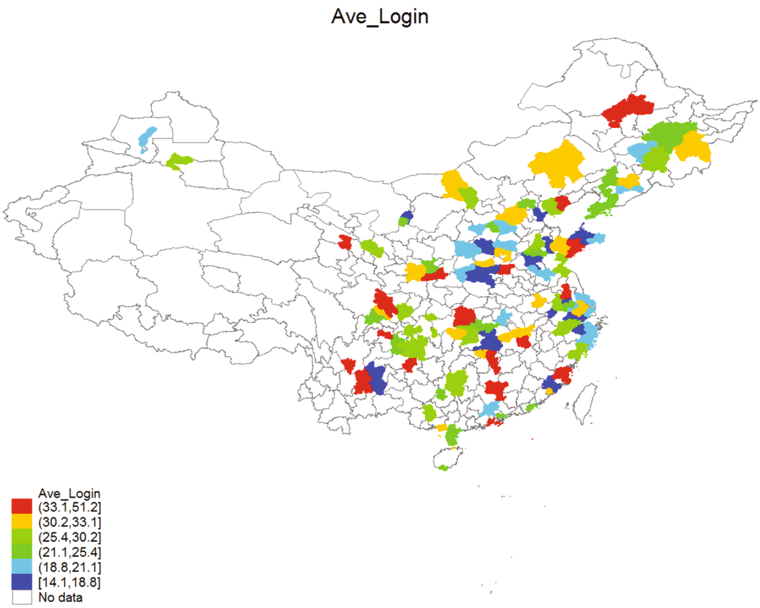


Fig. 1. Spatial distribution of average login time across 108 prefectures in China.

of playing cyber games are higher compared to low income area. In addition, metropolitans have more options for leisure activities such as concerts, shows, etc. It can be seen from column (5), which is our main model, that the higher numbers of movie theaters per person is correlated with less game time. The marginal effect of alternative leisure activities is very robust, with the coefficient being very close in magnitude and significantly negative as shown in columns (3)–(7). This shows the substitution effect of other leisure activities is strong for the average playing time. Notice that the income level is highly correlated with the number of movie theaters in our data, as shown in column (3) and (4), such that the coefficient of income variable is not significant if we only use movie as control. Unemployment rate is positively related to the playing time after controlling other variables. As the unemployment rate increase by 1%, the average play time increase by 0.06 h. This again supports the theory of time allocation: in aggregate level, the time spent on online games is negatively correlated with regular work. The environmental variables are not significant in the regression whether we use rainfall or air pollution variables. This environmental variable result shows that on the average level playing behavior is not affected by the outside environment.

Next, we turn attention to the achievement variable. Table 4 shows the regression results. Most importantly, the number of missions accomplished in the game is also negatively correlated with income level. An increase of 1% in income level, which is about 220 RMB for the sample average, is associated with the decrease of

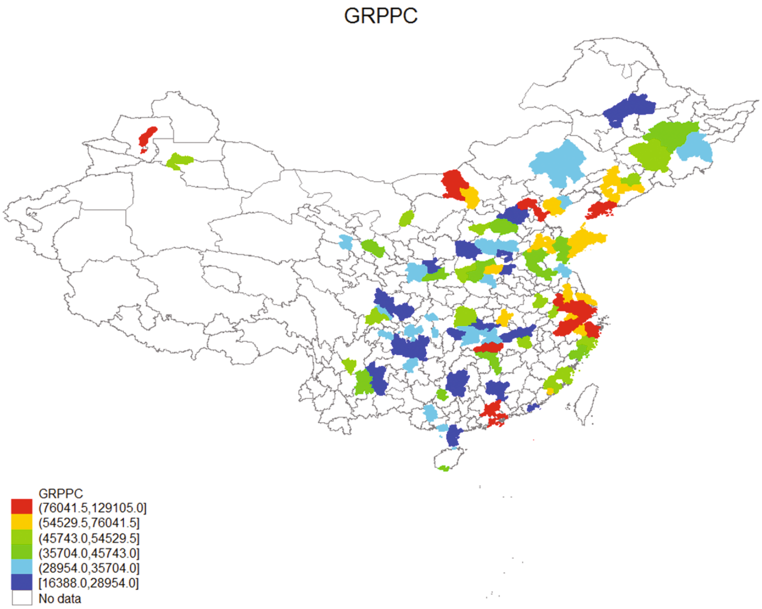


Fig. 2. Map of GRPPC of 108 prefectures in China. *Notes*: Figure shows a negative relationship between average playing time and GRPPC, this observation is formally studies in the following regression analysis.

Table 3. OLS estimation of login time

	Dependent variable: log average login time						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Log income	-0.3436^{***} (0.1005)		-0.2263^{*} (0.1294)		-0.4848^{***} (0.1793)		-0.3707^{**} (0.1671)
Log GRPPC		-0.1676^{***} (0.0507)		-0.1116^{*} (0.0636)		-0.2806^{**} (0.1079)	
Movie			-0.0644^{*} (0.0359)	-0.0664^{*} (0.0348)	-0.068^{*} (0.0354)	-0.0732^{**} (0.0339)	-0.0695^{*} (0.0364)
Int. penetration rate					0.5586^{**} (0.2564)	0.7357^{**} (0.2974)	0.4464^{*} (0.2504)
Unemployment					0.0612^{*} (0.0329)	0.068^{**} (0.0335)	0.0672^{*} (0.0344)
Precipitation					0.0013 (0.0009)		
Days $>$ 0.1 mm						0.0037 (0.0043)	
Log API							-0.061 (0.125)
Observations	108	108	108	108	108	108	108
R-squared	0.0747	0.0728	0.0994	0.1003	0.1657	0.1795	0.1501
p-value of F test	0.0009	0.0013	0.0002	0.0004	0.0001	0.0001	0.0001

Notes: Standard errors are reported in parentheses. Significance levels 0.1, 0.05 and 0.01 are noted by *, ** and *** respectively. Intercept are included but not reported.

Table 4. OLS estimation of mission

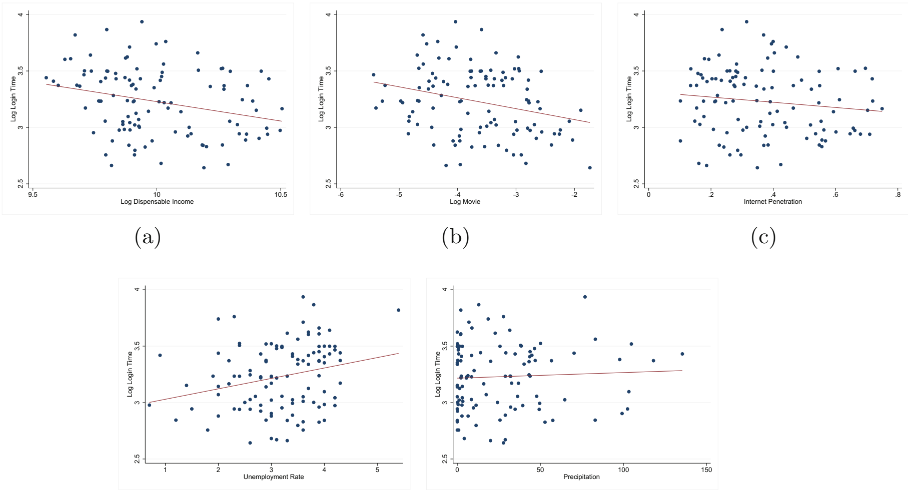
	Dependent variable: log average mission						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Log income	-0.0981^{***} (0.0349)		-0.0558 (0.0402)		-0.1612^{**} (0.0670)		-0.1336^{**} (0.0661)
Log GRPPC		-0.0406^{***} (0.0149)		-0.0197 (0.0186)		-0.0943^{***} (0.0331)	
Movie			-0.0230^{**} (0.0116)	-0.0253^{**} (0.0123)	-0.0261^{**} (0.0116)	-0.0284^{**} (0.0122)	-0.0268^{**} (0.0118)
Int. penetration rate					0.2574^{***} (0.0881)	0.3260^{***} (0.0970)	0.2302^{***} (0.0870)
Unemployment					0.0334^{***} (0.0103)	0.0364^{***} (0.0109)	0.0344^{***} (0.0103)
Precipitation					0.0003 (0.0003)		
Days > 0.1 mm						0.0004 v	
Log API							-0.0112 (0.0719)
Observations	113	113	113	113	113	113	113
R-squared	0.0489	0.0367	0.0758	0.0707	0.1950	0.2070	0.1870
p-value of F test	0.0058	0.0075	0.0023	0.0012	0.0001	0.0000	0.0001

Notes: Standard errors are reported in parentheses. Significance levels 0.1, 0.05 and 0.01 are noted by *, ** and *** respectively. Intercept are included but not reported.

0.10 of accomplished missions. As it shows in column (5), the coefficient increase in magnitude to 0.16 after we control for more variables. It suggests that the average performance, or game-playing 'productivity' is higher in lower income regions. Performance in cyber games is related to human capital, as the game-playing requires many skills that reflect human capital. However, the higher performance in cyber games does not translate to higher local economic development level. On the contrary, it shows the opposite association. This might infer the negative effect of playing cyber games in China, though we need to investigate more in the link between playing cyber games and income level. Internet penetration rate and unemployment are positively related to missions. Similar to the study of playing time, environmental variables are not significant here.

For level data, the OLS regression results are not very significant except for that of income and substitute effects. The underlying issue with the level data is that level does not reflect the quality of players in respect to playing skills. Also, some players who already achieved level 40 in our sample which is the maximum level one can get from playing this game. But the player can keep accomplishing new missions after achieving level 40. And even for players whose level is not yet 40, the level variables are accumulative and current level depend heavily on when the players starts to play the game. The starting time is unfortunately not recorded in our data. All relationships between the three LHS variable and RHS control variables are shown in Figs. 3, 4 and 5.

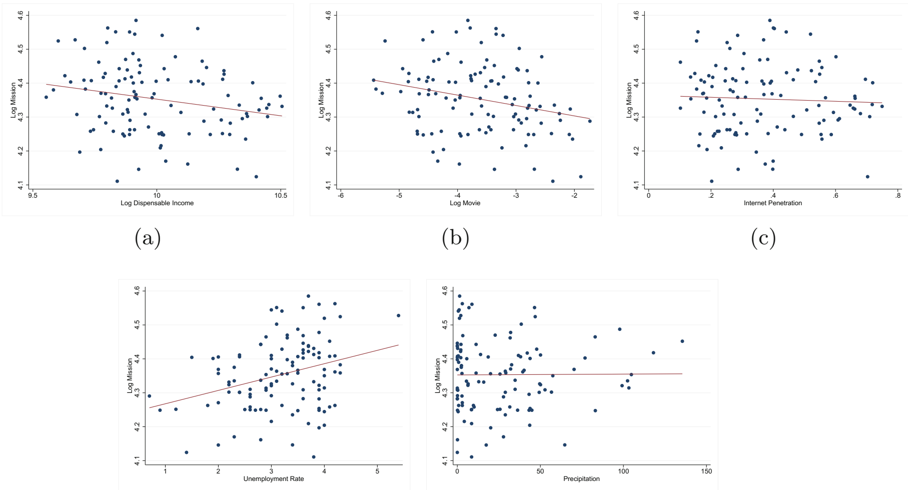
In summary of the OLS results, there are substantial evidence that after controlling for other determinant of cyber game participation, the coefficient of income is negative, which suggests that on prefecture-level, the effect of income



(a) (b) (c)

(d) (e)

Fig. 3. Scatter plot of login time and control variables



(a) (b) (c)

(d) (e)

Fig. 4. Scatter plot of achievements and control variables

on cyber game participation is negative. It could imply that cyber game participation significantly decreases with the increase in income level (Table 5).

The estimation on the effect of income on cyber game participation might be biased by the problem of reverse causality and omitted variables. To address the endogeneity concerns, we employ the instrumental variable freight volume which is exogenous to cyber game participation but is correlated with regional income

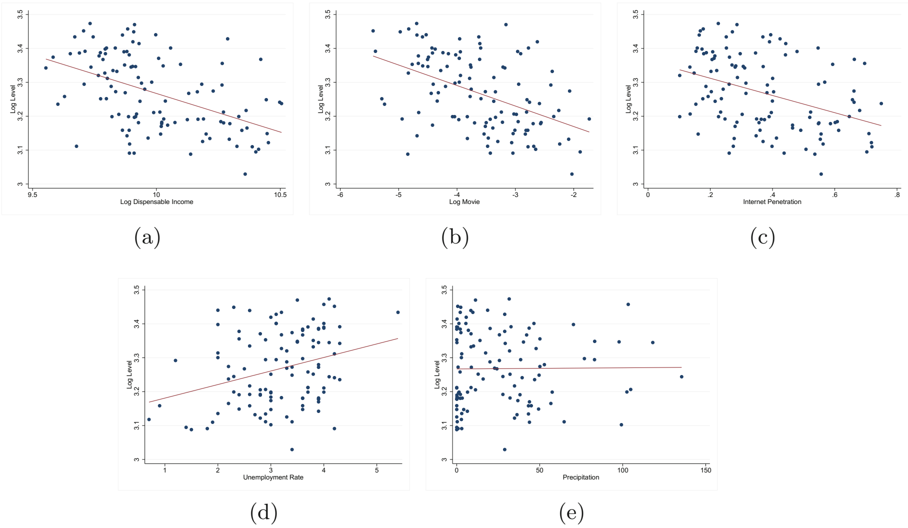


Fig. 5. Scatter plot of level and control variables

Table 5. OLS estimation of level

	Dependent variable: log average level						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Log income	-0.2275^{***} (0.0372)		-0.1552^{***} (0.0460)		-0.1677^{***} (0.0637)		-0.1309^{**} (0.0632)
Log GRPPC		-0.0981^{***} (0.0161)		-0.0613^{***} (0.0193)		-0.0567^{*} (0.0326)	
Movie			-0.0392^{***} (0.0135)	-0.0439^{***} (0.0131)	-0.0380^{***} (0.0134)	-0.0419^{***} (0.0134)	-0.0385^{***} (0.0136)
Int. penetration rate					0.0411 (0.0789)	0.0174 (0.0966)	0.0130 (0.0765)
Unemployment					0.0152 (0.0123)	0.0194 (0.0117)	0.0183 (0.0129)
Precipitation					0.0003 (0.0003)		
Days > 0.1 mm						0.0004 (0.0004)	
Log API							−0.0545 (0.0442)
Observations	112	112	112	112	112	112	112
R-squared	0.2334	0.1884	0.3027	0.2784	0.3195	0.2963	0.3216
p-value of F test	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

Notes: Standard errors are reported in parentheses. Significance levels 0.1, 0.05 and 0.01 are noted by *, ** and *** respectively. Intercept are included but not reported.

level. Our findings are mostly robust as shown in Table 6 using instrumental variable method. Specifically, income is significantly negatively associated with playing time. Substitutes, Internet access and unemployment all have significant and the signs agree with previous OLS findings. Similar results are also for game

Table 6. 2SLS estimation on impact of cyber game attributes

	Login time		Mission		Level	
	(1)	(2)	(3)	(4)	(5)	(6)
Log income	−0.2142** (0.1060)		−0.3956* (0.2235)		−0.2408 (0.1733)	
Log GRPPC		−0.5299*** (0.2040)		−0.1225 (0.0822)		−0.0701 (0.0811)
Movie	0.0015 (0.0024)	−0.0029** (0.0012)	−0.0011* (0.0006)	−0.0010** (0.0004)	−0.0003 (0.0006)	−0.0008** (0.0004)
Int. penetration rate	0.2033** (0.1013)	0.1299** (0.0517)	0.3786* (0.2648)	0.3786* (0.2081)	0.0452 (0.1555)	−0.0027 (0.0254)
Unemployment	0.0137 (0.0547)	0.0580* (0.0351)	0.0274* (0.0151)	0.0360*** (0.0117)	0.0156 (0.0145)	0.0210 (0.0121)
Precipitation	0.0041* (0.0021)	0.0006 (0.0008)	0.0001 (0.0005)	0.0004 (0.0001)	0.0001 (0.0001)	0.0001 (0.0001)
Observations	108	108	113	113	112	112
R-squared	0.0807	0.1585	0.0894	0.2016	0.2595	0.2428

Notes: Standard errors are reported in parentheses. Significance levels 0.1, 0.05 and 0.01 are noted by *, ** and *** respectively. Intercept are included but not reported.

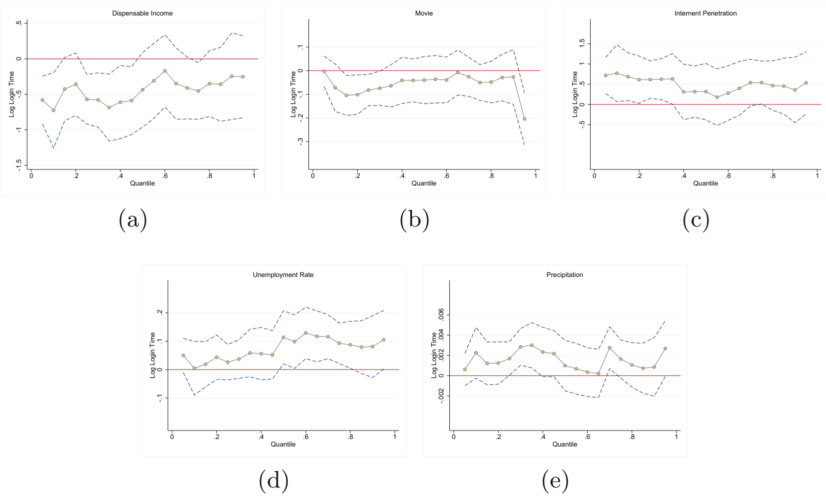


Fig. 6. Quantile regression estimates: login time. *Notes*: Figure graphs the estimated effect of income, substitute goods, etc., on the log of average login time at each quantile of the conditional distribution of the log of average login time and the associated 95% confidence interval.

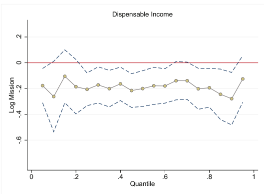
performance measured by missions. The results are not significant for levels partly due to the aforementioned reasoning.

For the quantile regression results, Table 7 shows that the income is negative and significant in lower quantiles of cyber game participation after controlling for other variables. The result is not significant for higher quantiles as seen in Fig. 6. This means that the effect of income is not very significant for the higher quantile of participation in the perspective of login time. And also that for lower

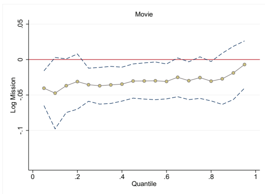
Table 7. Quantile regression results for login time

	Quantiles				
	5th	25th	50th	75th	95th
Log income	−0.5801*** (0.2056)	−0.5705*** (0.2124)	−0.4353 (0.3199)	−0.453* (0.2425)	−0.2526 (0.3522)
Movie	−0.0023 (0.039)	−0.082** (0.0403)	−0.0399 (0.0606)	−0.0501 (0.046)	−0.2039*** (0.0668)
Int. penetration rate	0.7156*** (0.2721)	0.6129** (0.2811)	0.3158 (0.4234)	0.5397* (0.3209)	0.5343 (0.4662)
Unemployment	0.0499 (0.0367)	0.026 (0.0379)	0.1138** (0.057)	0.093** (0.0432)	0.1052* (0.0628)
Precipitation	0.0006 (0.001)	0.0017* (0.001)	0.001 (0.0015)	0.0016 (0.0011)	0.0027 (0.0017)
Observations	108	108	108	108	108
Pseudo R^2	0.1331	0.1086	0.0965	0.0684	0.1415

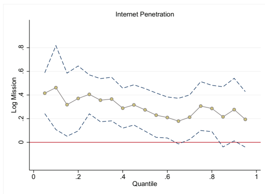
Notes: Standard errors are reported in parentheses. Significance levels 0.1, 0.05 and 0.01 are noted by *, ** and *** respectively. Intercept are included but not reported.



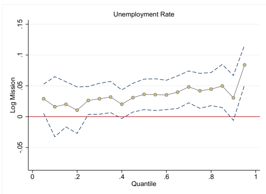
(a)



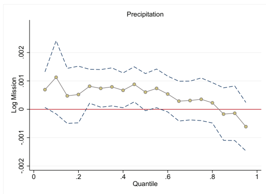
(b)



(c)



(d)



(e)

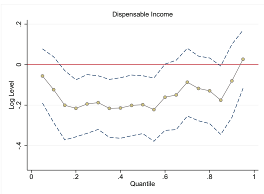
Fig. 7. Quantile regression estimates: mission. *Notes*: Figure graphs the estimated effect of income, substitute goods, etc., on the log of average login time at each quantile of the conditional distribution of the log of average login time and the associated 95% confidence interval.

quantiles, the absolute value coefficient of income is significantly higher than the higher quantiles. Therefore players in the lower quantiles of playing cyber games is more responsive to income level. We find it interesting that substitute activities such as movies are significant in lower quantile, but not so in higher quantile. Furthermore, for the highest quantile, it is significant again. Given the small sample size in the highest quantile, this result could be affect by a

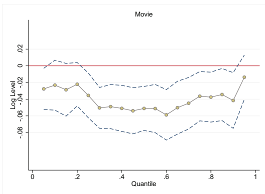
Table 8. Quantile regression results for mission

	Quantiles				
	5th	25th	50th	75th	95th
Log income	−0.1772** (0.0808)	−0.2058*** (0.0770)	−0.2002** (0.0841)	−0.2018** (0.0967)	−0.1252 (0.1098)
Movie	−0.0404*** (0.0149)	−0.0355** (0.0142)	−0.0302* (0.0155)	−0.0255 (0.0178)	−0.0070 (0.0203)
Int. penetration rate	0.4154*** (0.1052)	0.4054*** (0.1001)	0.2743** (0.1095)	0.3063** (0.1257)	0.1938 (0.1428)
Unemployment	0.0291** (0.0144)	0.0263* (0.0137)	0.0362** (0.0150)	0.0419** (0.0173)	0.0839*** (0.0196)
Precipitation	0.0007* (0.0004)	0.0008** (0.0004)	0.0006 (0.0004)	0.0004 (0.0005)	−0.0006 (0.0005)
Observations	113	113	113	113	113
Pseudo R^2	0.1496	0.1222	0.1346	0.1359	0.1527

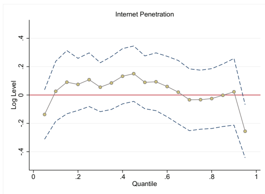
Notes: Standard errors are reported in parentheses. Significance levels 0.1, 0.05 and 0.01 are noted by *, ** and *** respectively. Intercept are included but not reported.



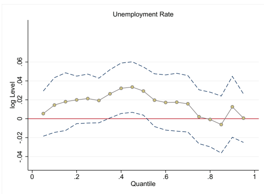
(a)



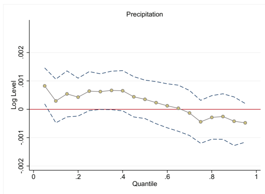
(b)



(c)



(d)



(e)

Fig. 8. Quantile regression estimates: level *Notes*: Figure graphs the estimated effect of income, substitute goods, etc., on the log of average login time at each quantile of the conditional distribution of the log of average login time and the associated 95% confidence interval.

few dominant regions. Or it could imply that for the very high quantile, movie or other substitute goods are becoming relevant again due to the fact that the game eventually lose its attractiveness to its top players which is very common in most cyber games. The cyber game company usually promote new games every two or three years to accommodate the decline of interests of its players. For lower quantile of cyber game participation, the Internet penetration rate is also

Table 9. Quantile regression results for level

	Quantiles				
	5th	25th	50th	75th	95th
Log income	−0.0560 (0.0813)	−0.1942** (0.0882)	−0.19772** (0.0868)	−0.1176 (0.0970)	0.0266 (0.0873)
Movie	−0.0277* (0.0150)	−0.0355** (0.0163)	−0.05112*** (0.0160)	−0.0365** (0.0179)	−0.0137 (0.0161)
Int. penetration rate	−0.1373 (0.1060)	0.1080 (0.1150)	0.0894 (0.1132)	−0.0329 (0.1265)	−0.2555** (0.1139)
Unemployment	0.0054 (0.0145)	0.0213 (0.0157)	0.02932* (0.0155)	0.0021 (0.0173)	0.0006 (0.0156)
Precipitation	0.0008** (0.0004)	0.0006 (0.0004)	0.0004 (0.0004)	−0.0004 (0.0005)	−0.0005 (0.0004)
Observations	112	112	112	112	112
Pseudo R2	0.1731	0.1819	0.2404	0.1842	0.1152

Notes: Standard errors are reported in parentheses. Significance levels 0.1, 0.05 and 0.01 are noted by *, ** and *** respectively. Intercept are included but not reported.

significant. Unemployment rate has significant influence on the higher quantiles. Weather variable is in general not significant except for some lower quantiles. This means the rainfall e.g., would have some impact on the infrequent players.

The quantile results of achievement is similar to the finding of login time as shown in Table 8. All control variables have significant impact on the lower quantile and most are insignificant for the very high quantiles. We find that income is mostly negative and significant as shown in Fig. 8(a). Substitute goods is significantly negatively correlated to achievements in lower quantiles but not in higher quantiles. Weather variable is showing some significantly positive correlation in the lower quantiles also.

As the reasons stated in OLS regression results, level variable has some problems in being a good proxy for game performance. We could still observe the income and substitute has significant impact for lower quantiles. Other control variables are not significant for most quantiles with few exceptions (Table 9).

5 Conclusion

This paper uses the gaming data and economic data to empirically analyze the participation of cyber games. Using a prefecture-level Chinese data in 2011, our empirical results show that the income level is negatively associated with game participation, while other economic variables and environmental variables have more impact on the lower quantiles. The substitution effect of other leisure activities are strongly negative to the participation of cyber games. We check the robustness of the results using different regression models and measurements of participation including playing time, level and achievements. The study of individual level data and the weather's impact on the individual player's performance would be an interesting future research topic.

Acknowledgments. This work is supported by the National Natural Science Foundation of China (No. 71671149, 71631004 (Key Project), 71301133, 71572166, and 71202059), Natural Science Foundation of Fujian Province of China (No. 2016J01340) and the Fundamental Research Funds for the Central Universities (Project No. 20720161044).

References

1. Auld, M.C., Grootendorst, P.: An empirical analysis of milk addiction. J. Health Econ. **23**, 1117–1133 (2004)
2. Baltagi, B.H., Griffin, J.M.: The econometrics of rational addiction. J. Bus. Econ. Stat. **19**(4), 449–454 (2001)
3. Becker, G.S.: A theory of the allocation of time. Econ. J. **75**(299), 493–517 (1965)
4. Becker, G.S., Grossman, M., Murphy, K.M.: An empirical analysis of cigarette addiction. Am. Econ. Rev. **84**(3), 396–418 (1994)
5. CNNIC: 2013 Chinese online gamer behaviour analysis report (2013). http://www.cnnic.cn/hlwfzyj/hlwxzbg/wybg/201402/P020140228346868241995.pdf
6. CNNIC: The 35^{th} china internet development statistical report (by Jan 2015) (2014). http://www.cnnic.cn/hlwfzyj/hlwxzbg/hlwtjbg/201502/P020150203548852631921.pdf
7. Debeauvais, T., Nardi, B., Schiano, D.J., Ducheneaut, N., Yee, N.: If you build it they might stay: retention mechanisms in world ofwarcraft. In: Proceeding of the 6th International Conference on the Foundations of Digital Games (FDG 2011), pp. 1–8 (2011)
8. Manning, W.G., Blumberg, L., Moulton, L.H.: The demand for alcohol: the dierential response to price. J. Health Econ. **14**, 123–148 (1995)
9. Olekalns, N., Bardsley, P.: Rational addiction to caffeine: an analysis of coffee consumption. J. Polit. Econ. **104**(5), 1100–1104 (1996)