CHARLES UNIVERSITY

FACULTY OF SOCIAL SCIENCES

Institute of Economic Studies

Bachelor thesis

2021 **Tomáš Karhánek**

CHARLES UNIVERSITY

FACULTY OF SOCIAL SCIENCES

Institute of Economic Studies

Tomáš Karhánek

Lottery premium in video gaming environment

Bachelor thesis

Prague 2021

Univerzita Karlova

Fakulta sociálních věd

Institut ekonomických studií

Tomáš Karhánek

Loterijní prémie v prostředí videoher

Bakalářská práce

Praha 2021

Declaration of Authorship

I hereby declare that I compiled this thesis independently, using only the listed resources and literature, and the thesis has not been used to obtain any other academic title.

I grant to Charles University permission to reproduce and to distribute copies of this thesis in whole or in part and agree with the thesis being used for study and scientific purposes.

Prague, February 11th 2020

Tomáš Karhánek

Acknowledgments

I would like to thank my supervisor, Mgr. Michal Šoltés, for his time, patience, guidance and valuable suggestions for the thesis. I would also like to thank video gaming community that helped me gather valuable data for this thesis. Last but not least, I would like to thank my family and friends that were a great support during my studies.

Bibliographic note

KARHÁNEK, Tomáš. Lottery premium in video gaming environment. Bachelor's thesis. Charles University, Faculty of Social Sciences, Institute of Economic Studies. Supervisor Mgr. Michal Šoltés.

Abstract

Gambling, as well as many traditionally physical forms of entertainment, has become popular in a digital form. In the recent years it has found its way to video games. In the environment the gambling element is represented by so called loot box, virtual vessel containing random digital rewards that are in most cases purchasable strictly for real world currency. The writer hypothesizes an existence of a variable, lottery premium, that can help him well explain the decision making process of gambling participants in this environment. The author collects data from the loot box buyers to help him determine the other significant factors influencing the amount of money spent on loot boxes. Using OLS regression he estimates the size and direction of effects of mentioned variables as well as their significance. The lottery premium variable is determined as a significant in terms of the amount spent. Therefore, the author performs a second OLS regression to determine factors influencing size of the lottery premium. The significant variables influencing the amount spent are: age, participation in risk involving activities, searching for the odds of the loot boxes, amount of friends that purchase loot boxes as well, the value of the most expensive item ever received from the loot box, disposable income and already mentioned lottery premium. The variables significantly influencing the size of lottery premium are: a type of risk attitude the individual expresses when purchasing loot boxes, education level, if the person watches the loot box openings and if the person purchases loot boxes at least once a month. The author presents a suggestion for a future research in the field.

Keywords Lottery premium, gambling, utility, loot boxes, OLS, video game industry, decision making process under uncertainty, risk attitude

Abstrakt

Hazard, stejně jako další typicky fyzické formy zábavy, se stal populárním i v digitální formě. V posledních letech si našel cestu do video herního průmyslu. V tomto prostředí je prvek hazardu reprezentován takzvaným loot boxem, to je virtuální nádoba obsahující digitální předměty, která se dá převážně pořídit jen za reálné peníze. Autor se domnívá, že existuje proměnná, loterijní prémie, která může pomoct vysvětlit proces, kterým se lidé při nákupu loot boxů řídí. Autor sbírá data od nakupujících, která mu pomůžou zjistit, jaké proměnné ovlivňují množství peněz, které lidé za loot boxy utratí měsíčně. Je použita metoda nejmenších čtverců, která pomůže odhadnout sílu a směr efektu, které jednotlivé proměnné mají, stejně jako signifikanci těchto proměnných v této regresi. Loterijní prémie je detekována jako jedna ze signifikantních a autor se rozhodl provést druhou regresi stejnou metodou, aby zjistil, které faktory ovlivňují velikost této prémie. Signifikantní proměnné, které ovlivňují množství peněz utracených měsíčně, jsou: věk, účast v aktivitách zahrnujících finanční riziko, zkoumání pravděpodobností loot boxů, množství přátel, kteří kupují loot boxy, hodnota nejdražší věci, kterou jedinec obdržel z loot boxu, disponibilní příjem a již zmíněná loterijní prémie. Proměnné signifikantně ovlivňující velikost loterijní prémie jsou: postoj k riziku, které nakupující zaujímá, vzdělání, zda respondent sleduje otevírání loot boxů na internetu a zda pořizuje loot boxy, alespoň jednou měsíčně. Autor předkládá návrh pro budoucí výzkum v této oblasti.

Klíčová slova Loterijní prémie, gambling, užitek, loot box, metoda nejmenších čtverců, videoherní průmysl, rozhodování v nejistotě, postoj k riziku

Bachelor's Thesis Proposal

Institute of Economic Studies
Faculty of Social Sciences
Charles University in Prague

Author's name and surname: Tomáš Karhánek

Supervisor's name: Mgr. Michal Šoltés

Notes: Please enter the information from the proposal to the Student Information System (SIS) and submit the proposal signed by yourself and by the supervisor to the Academic Director ("garant") of the undergraduate program.

Proposed Topic:

Lottery premium in video gaming environment.

Preliminary scope of work:

Research question and motivation

The main research question of this thesis is to test whether risk-seeking behaviour of customers is sufficient to justify participating in video-gaming industry gambling, when an alternative of purchasing same reward in a risk-free way for fixed price exists. Alternatively, I hypothesize presence of lottery premium suggesting that consumers derive utility not only from the lottery outcome itself, but also from playing the lottery. I suppose that video-gaming environment; in which subjects (arguably) derive utility from playing is a proper environment to study such behaviour. Understanding if and when players pay to play may help us understand why some individuals spend significant amount of money playing lottery despite a negligible chance of winning and potentially mimic negative externalities and internalities of their behaviour. .

Loot boxes (loot crates) are nowadays known as virtual consumable items generating randomized virtual content of different value (quality), often purchasable only via real money. A concept of loot boxes (known since 1800s) experienced great boom in the 90s with CCG[1] (e.g. basketball cards, Magic: the Gathering and Pokémon). Some of those yielded revenue over 1 billion USD per year (Jamieson 2011). An implementation into video game industry began in 2004, but a real expansion across all platforms and games happened in 2016. Earnings from loot crates are 30 billion USD worldwide per year. It is estimated that by 2022 the amount will be 50 billion USD per year and further rise in future (Juniper Research 2018).

A controversy of the loot boxes is that some authorities consider them to be a form of gambling aimed at children and adolescents, therefore it lies in a grey zone between gambling and gaming.

The speciality of gambling in this environment is that in most cases it is not driven by maximising wealth in terms of money and that the reward from the gambling is often purchasable through risk-free transaction for fixed cost. The incentive of investment is unlikely to be important, as the inner market (if present) does not provide extraction of money from the system in official way (money laundering prevention) and the existing black markets, where products could be potentially monetized, are characterized by weak enforcement of contracts. Furthermore, the amount of money spent on loot boxes is higher than market value of the items obtained from them (based on expected value of loot box). Therefore, purchase of loot boxes cannot be considered generally as investment opportunity. People selling virtual items under price of initial investment often justify their behaviour as getting back part of the sunk costs.

My hypothesis states that in this market exists a variable representing the lottery premium which causes that the customers are willing to participate in the gambling although the expected value is lower, even after a correction for the risk-seeking behaviour, than value of the risk-free purchase of a specific item and that the

[1] Collectible card games, the cards were on sale in form of boosters that contained random cards.

variable is significant with respect to the price that people are willing to pay for the loot boxes. In other words, the players are willing to pay more money for the random process with an unsure result than for the same sure one. Once shown that this premium exists, I plan to study significant factors influencing the volume of the premium. Once we understand the motivation behind participating/not participating in gambling in the market we can use the findings to elicit decision-making process of the individuals in this environment. The importance of this research lies in providing us with essential information to perform further studies, namely impact evaluation focused on regulations levied on loot boxes and finding the optimal one. Furthermore, despite relatively low external validity, my results may help us to understand gambling behaviour.

Contribution

Although governments pass laws that radically influence this environment (bans, age restrictions) there has been no economic evaluation written on this specific market so far. There exist several psychological papers such as: Is the buying of loot boxes in videogames a form of gambling or gaming?, written by Professor Mark D. Griffiths, but from the perspective of economics this issue remains underexplored. I aim to contribute to behavioural economic research at the field of gambling and extend it to video-gaming industry, which is going through a transformation of shifting used to be model of pay-to-play into pay-to-win[2] scheme. A development in virtual reality promises a great future to the industry, whose market value is estimated to growth from 135 billion USD value today up to 160 billion USD by 2022 (Juniper research 2018). To be able to create valid economic predictions and models that can help find the middle ground among regulators, distributors, and customers we must understand the risk preferences and other motivators of gambling participation of the customers, therefore it is important to develop economic research at this field and its monetization now.

Methodology

At first I will use already existing knowledge on decision-making process, risk preference (Gul 1991, Kwang 1965), utility in gambling (Conlisk 1993), and data collected through a survey to develop a utility function capturing the risk aversion/risk seeking behaviour (O´Donoghue and Somerville 2018). From there I will be able to create a linear regression model. Using OLS method and with the knowledge of willingness to pay (from the survey) I will be able to firstly estimate lottery premium and confirm whether it is significant variable regarding the willingness to pay or not and secondly the influence of other variables in terms of the lottery premium volume. Variables to be used in lottery premium estimation shall be e.g. age, monthly income, monthly spending on loot boxes, time played per week, etc.

As there is a game heterogeneity (existence of inner market, ability to purchase the specific good directly without risk, quantity of monthly active players, initial purchase of game, etc.), whose omission would cause bias at game level analysis, in which they can be replaced by a fixed effects model. If the lottery premium appears to be significantly higher in a specific game, we can use so far omitted variables to help us understand why it is so.

I will work with the data collected through the survey that will help me understand the decision-making process at the customer level. I will be performing survey during 8 months period starting in August 2019 and finishing in March 2020 on monthly basis through different communities of players on live streams, forums, fan pages and YouTube channels and game channels to gain data from as heterogeneous sample as possible to validly represent the whole population of players. To cover most often types of loot boxes I will gather the data on the 4-6 most popular games from which each one represents a specific type of loot box.

The data that I will collect consist of: age, sex, nationality, education (including current study), time played per week, monthly income, source of income spent in-game on loot boxes (monthly), the way of purchase (credit card, Paysafecard, Pay Pal), purchases on regular basis or increased during special events, number of friends playing with the subjects regularly (at least once a week), watching streaming platforms featuring the game (YouTube/twitch.tv) and more. To elicit willingness to pay in systems without inner market, I will include

[2] The scheme based on monetization through microtransactions among which loot boxes belong.

questions regarding the price for which they would sell their unopened loot box.

Outline

Abstract
Introduction
a) Brief history and current situation
b) Specifications of the market
c) Specification of research questions
Literature review
a) Literature on utility in gambling environment
b) Literature on risk preferences
c) Knowledge so far on this subject
d) Specifying the hypothesis to be tested
e) Justifying the hypothesis
Methodology
a) Description of data
b) Relevance of data
c) Utility function
d) Willingness to pay elicitation
e) OLS
Results
a) Rejecting/not rejecting the hypothesis
b) Showing significance/ no significance
c) Justifying different results (if exist) for different games
Conclusion
a) Interpretation of results and their impact
b) Further research

List of academic literature:

Bibliography

C. Faustino, Horácio & João Kaiseler, Maria & Marques, Rafael. (2009). Why Do People Buy Lottery Products?.

Chateauneuf, A. & Cohen, M. J Risk Uncertainty (1994) 9: 77. https://doi.org/10.1007/BF01073404

Thaler, Richard & Johnson, Eric. (1990). Gambling With the House Money and Trying to Break Even: The Effects of Prior Outcomes on Risky Choice. Management Science. 36. 643-660. 10.1287/mnsc.36.6.643.

Ng, Yew-Kwang. (1965). Why do People Buy Lottery Tickets? Choices Involving Risk and the Indivisibility of Expenditure. Journal of Political Economy. 73. 530-530. 10.1086/259077.

Conlisk, J. J Risk Uncertainty (1993) 6: 255. https://doi.org/10.1007/BF01072614

Gul, Faruk. “A Theory of Disappointment Aversion.” *Econometrica*, vol. 59, no. 3, 1991, pp. 667–686. *JSTOR*, www.jstor.org/stable/2938223.

O'Donoghue, Ted, and Jason Somerville. 2018. "Modeling Risk Aversion in Economics." *Journal of Economic Perspectives*, 32 (2): 91-114, https://pubs.aeaweb.org/doi/pdf/10.1257/jep.32.2.91.

Contents

List of tables and figures

Chapter 1 Introduction

Gambling, as a cultural phenomenon, is inseparable part of a human behaviour as there were and always will be people willing to undergo a risk with a vision of fortune or how Friedman and Savage (1948) stated in their paper: “young men of an adventurous disposition”.

The term gambling often invokes a negative associations and risk of losing money, while gambling is essentially a participation in any type of action involving money that includes multiple possible outcomes regardless of the associated odds. Throughout the history the games of chance have evolved, from dices to card games of all kind, from cards to roulettes, from roulettes to slot machines, but they always had a certain aspects in common. It was officially declared as gambling activity, which means that the officials knew it was gambling as well as the participants. The authorities could pass laws to set clear rules and possible taxations related to it.

The taxation is another factor influencing the gambling. It became convenient for the governments to keep official gambling places, such as casinos, open and benefit from the taxation of their profits. Profits of the providers of gambling services are in the most states taxed with an excise tax, whose rate is similar to excise tax for alcohol and tobacco products.[3]The USA generated 41 billion USD[4] in 2019 just as tax revenue from the taxing of commercial gambling activities. Consequently, the extensive gambling, as well as already mentioned alcohol and tobacco products, might be cause of health issues. Even though, the health issues associated with gambling are mostly related to mental health. The WHO officially recognizes gambling addiction as a disease[5] and in 2019 there were estimated 2 million citizens suffering from the gambling addiction in the USA[6].

From the point of view of behavioural economics there was a significant amount of experiments and papers worked out on the gambling in general, but no single rigid theory persists in all possible circumstances. The author chose gambling in an environment of video games where as a gambling element serve so called loot boxes. The speciality of this particular field of gambling and simultaneously the reason why the writer chose it lies in several aspects.

The most important one is because the whole process is on the virtual level. The gambling process is virtual and the possible rewards associated with it are virtual as well. The other property that distinguishes this gambling from others is that the rewards do not consist of

[3] Czech Republic has increased excise tax for lottery games to 35% and cigarettes are taxed at 30% rate + fixed amount per cigarette (2.9 CZK) as of 2020. Available at: https://www.finance.cz/530435-spotrebni-dan-tabak-alkohol-zdaneni-vyher/#3

[4] https://www.statista.com/statistics/187972/number-of-us-commercial-casinos-since-2005/

[5] https://sudonull.com/post/14890-World-Health-Organization-officially-recognized-the-existence-of-gambling-addiction

[6] https://www.psycom.net/drug-alcohol-addiction-treatment/gambling-addiction/

direct money gains. Last but not least, the fact that this particular gambling is accessible to adolescents and children is what troubles the regulators.

The concept of gambling in this environment is nothing new; it works on the same principle as packages of random cards for a card gathering game (e.g. Pokémon), but its implementation into the most of the video-games in recent years caused a large spike in the overall industry earnings.[7] As a response to the rising issue, several countries, their regulators respectively, passed laws that had more or less severe impacts on the business model of the game distributors, depending on the country. Besides the unique aspects of the market, there are subsequent properties that might bring attention of more economists to perform research. The market is new and rapidly growing. In 2017 earnings from the boxes alone were around 30 billion USD with prediction to grow up to 50 billion USD by 2022 driving the market total worth to 160 billion USD by the same year[8]. Furthermore, with the progressive digitalization, the industry reported over 90% of the revenue coming from the digital trade in 2020.[9]

Loot boxes in games existed, unnoticed by the regulators, in the grey zone between gambling and regular microtransactions for several years. Since the regular law structure of gambling restrictions was not able to reach them, nobody could control who is actually purchasing the boxes and unlike slot machines in casinos that have to have officially stated amount of wins in a certain finite number of playing cycles, there is no legislative structure that would force the loot box sellers to state the real odds on the individual items. The process of implementing restrictions on the loot box purchasing process is also getting more complicated by the fact that there are several purchasing methods. Based on the research performed in 2018 by ACI Worldwide and Newzoo, both companies operating in the gaming industry, it was found out that the most prevailing payment method is PayPal.[10]

The existence of several ways of payment results in the fact that it is almost impossible to restrain certain group of people from the gambling participation. Therefore, some governments decided to pass laws affecting their entire population. One of such countries is Belgium which recognized loot boxes to be subject to the gambling law in the country and since the separation of the non-adults and adults was impossible, the authority stated that loot boxes must be removed from all the games distributed in Belgium under the threat of 800,000 EUR fine and up to 5 years in prison for the publisher. This extension of the law has been in effect since 2018.[11]

The writer's goal in this thesis is to find out which theory of utility in gambling environment supports the results of the thesis and use it to explain certain phenomenon that might arise along the way. The author hypothesizes that there is a variable influencing decision making process which is related to the fact that the potential utility from the gambling came from the

[7] Source: https://www.wepc.com/news/video-game-statistics/
[8] Source: https://www.juniperresearch.com/document-library/white-papers/in-game-gambling-the-next-cash-cow
[9] Source: https://flipboard.com/@GameRant/91-of-2020-s-game-industry-revenue-was-digital/a-WmsScqIeTh6TJ1lbUoJggQ%3Aa%3A827674793-4ae3d594d4%2Fgamerant.com
[10] Source: https://newzoo.com/insights/articles/aci-and-newzoo-whitepaper-turning-players-into-payers/
[11] Source: https://www.bbc.com/news/technology-43906306

process of gambling itself, therefore building up upon the theory of Conlisk (1993), who stated in his paper that besides the utility the individual can obtain from gambling participation there is also a utility strictly related to the gambling process itself. Hence, the gambling in this environment transforms into a purchase of digital assets with uncertain outcome. The author calls this variable the lottery premium. It represents the utility that the individual gets from undergoing the excitement of unsure purchase of the loot box even though it yields the same result as risk-free purchase of the same goods. The author is aware of the risk premium existence.

To understand the difference between both premiums author defines and compares both. Risk premium (RP) is commonly related to financial and investment market where it represents the difference between expected return on risky asset (e.g. stocks, shares) and risk-free alternative (risk-free bond[12]). However, a risk premium, as mentioned, is not bound to financial markets solely. In general terms it represents a quantified difference in utility (usually expressed in terms of money) between expected utility of a risk involving choice (purchasing a lottery ticket) and risk free choice (cost of the lottery ticket). Thus, RP is dependent on the expected utility (value) of undergoing risk involving investment and individual risk preferences in general.

Lottery premium (LP) is the quantified utility resulting directly from the participation in the lottery (the joy from the game itself regardless of the result, the loot box purchase and opening). The LP in this thesis represents the difference between option involving certain gain of money, gain of money involving chance and willingness to pay for loot box purchase with the same odds and expected values as the monetary gain options. Therefore, LP is indirectly influenced by the difference of a risk attitude that the individuals are expressing based on the presence of the game factor.

The main aim of the thesis is to show that expected utility theory itself is not valid in this case and prove if this industry exists the mentioned lottery premium and whether it represents a significant variable in decision making process of purchasing loot boxes or not. If that is the case, the author will then estimate impact of variables of the individuals participating in the lottery (e.g. age, education). If the LP proves to be significant factor, he will test variables, influencing the lottery premium itself, for significance.

The possible contribution of the thesis lies in creating a framework for subsequent economic researches in this area as well as for authorities being responsible for regulating possible gambling opportunities, providing both with information on customers´ decision making and significant factors influencing it so they are able to introduce more effective (not excessive) restrictions and regulations and perform further research.

[12] From the nature of the real world economy there are no perfectly risk-free assets, for our purpose here we assume a theoretical bond that yields the same amount regardless of conditions.

Chapter 2 Literature review

This chapter is dedicated to reviewing and analysing the already existing literature on the decision making process under uncertain outcomes and underlying utility of the individuals. To the author's knowledge there is no piece of behavioural economics literature written on this specific topic that the author could use. Hence, the author also analyses the academic papers and properties of the environments in which they were written to find analogy to the video-game market. The amount of literature written on the decision making process under uncertain circumstances and related utility of the decision makers is vast. Thus, the author decided to follow main line of economical thought with a few extra papers that might be helpful in understanding the decision making process.

2.1 Utility of gambling

The main stream of the papers dealing with utility of gambling begins at 1738 with very well-known St. Petersburg Paradox. Though the solution and theory was published by Daniel Bernoulli (everyone has already encountered this gentleman during physics classes about fluid mechanics in high school) in 1738 it was in fact his cousin – Nicolas Bernoulli who invented the problem already in 1713. The author presents the problem and theory behind it very briefly as it is not suitable for the thesis except as the foundation stone upon which the rest of the papers are built. The paradox is a game of chance for a one player. The player is presented with a fair coin. The game begins with the stake of 2 dollars and in each stage the coin is tossed. If the heads appear new stage begins, the amount of money in the pool is doubled and the process repeats until the first tail appears. At that point the player wins all the money in the pool. The riddle of the paradox is: What is the fair price a player should pay casino to participate in this gamble? If one approaches the problem using expected utility, he/she will find himself/herself quickly trapped in a sum of non-finite series that is divergent as the expected value calculation for this problem looks as follows:

$$E = \frac{1}{2} * 2 + \frac{1}{4} * 4 + \frac{1}{8} * 8 + \frac{1}{16} * 16 + \frac{1}{32} * 32 + \cdots = 1 + 1 + 1 + 1 + 1 + \cdots = +\infty$$

To solve the problem, Bernoulli took the known utility properties of utility function and laid foundation of the expected utility hypothesis, presuming diminishing marginal utility. He then stated that one must not look at the problem as the wealth maximisation issue but rather as utility maximisation. To capture diminishing utility of money he described utility as logarithmic function of wealth:

$$U(w) = \ln(w),$$

where w represents total wealth of a player.

He used this concept and presented equation solving the problem.

$$\Delta E(U) = \sum_{k=1}^{+\infty} \frac{1}{2^k} [\ln(w + 2^k - c) - \ln(w)] < +\infty, k \in N; w, c \in R^+,$$

where k is a number of tosses and c represents the price of the gamble. 2^k is the mathematical yield of the gamble and 2^{-k} the probability of win.

The expected utility theorem has been further developed by several authors. The most significant benefit to expected utility hypothesis and the decision making theories based on Bernoulli' s work came from John von Neumann and Oskar Morgenstern (1947) and their common work - Theory of Games and Economic Behavior (1947) in the book they created what is nowadays known as VNM utility theorem. The VNM theorem describes well why risk seeking individuals are participating in the gambling as well as risk-averse individuals. It does so by not using the already set idea of people making decision based on the utility, but the state of wealth itself.

They incorporated four axioms that well define people's decisions that would be otherwise in contradiction with expected value conventions. The general idea of VNM is that under those stated axioms, an individual performing decision-making process involving risk/uncertain outcomes behaves in the way of maximization expected value of a certain function which is defined over the set of potential outcomes. In other words, every individual that satisfies all four VNM stated axioms has a utility function which he will always maximize. Thus, the individual can be considered (VNM) rational. The axioms mentioned in this section are:

1) Completeness – For any lotteries K,L, one of the following holds:
   $K \prec L$, $L \prec K$,or $L \sim K$
2) Transivity – For lotteries K,L,M; if $K \prec L$ and $L \prec M$, then $K \prec M$.
   This also hold for the indiference.
3) Continuity – For lotteries K,L,M; if K is weakly preferred to L and M is weakly preferred to L then there exists a probability $P \in [0,1]$ such that $PK + (1 - P)N \sim L$
4) Independence – For any M and $P \in (0,1]$,
   K is weakly preffered to L if $PK + (1 - P)M \preceq \mathrm{PL} + (1 - \mathrm{P})\mathrm{M}$

Consequently the VNM theorem became basis for expected utility theory as it is known today. Even though it underwent a significant improvement, the expected utility theory was still too limited to explain everyday situations where people simultaneously purchase lottery tickets and insurance. The hypothesis was too rigid with the relationship between the shape of utility function of wealth and the risk preferences. Put plainly, risk aversion individuals were associated with concave shape of the utility of the wealth function, which implies decreasing marginal utility of wealth. On the other hand, the risk seeking individuals were assumed to have convex shape of the said utility function, which represents increasing marginal utility of wealth.

The rigid shape of the utility function of wealth for either individual was challenged by Milton Friedman and Leonard J. Savage, who only the year after the prior duo came with VNM theorem, came with their paper – The utility analysis of choices involving risk (1948) explaining small gambles contradicting general nature of a decision maker. The mentioned authors came with an idea that the curve of utility function of income has a certain ranges depending mostly on the economical class, in other words – wealth, at which the utility function even for risk-averse individual changes from concave shape into convex for a certain levels of income, consequently creating a theoretical support for the risk-averse decision makers participating in gambles that would otherwise mean breach of rationality. On the other hand, it helps explain purchases of insurance for a risk-seeking individual as they happen to have concave parts of theirs utility of income functions. Friedman and Savage also test for the riskless alternative and which properties must the utility function of the income satisfy for a decision maker to choose it over a gamble alternative.

Figure 2.1: From the original paper published by Friedman and Savage

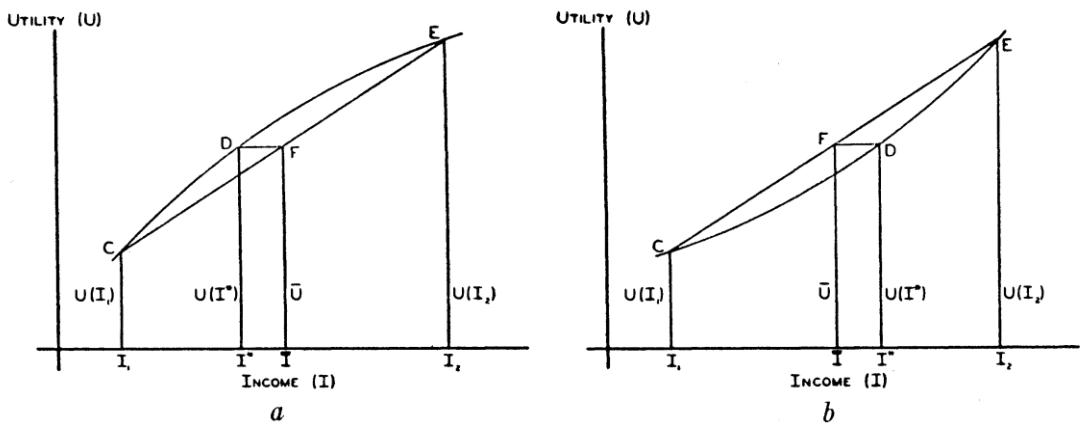


FIG. 1.—Illustration of utility analysis of choices involving risk: *a*, preference for certainty; *b*, preference for risk.

The figure 2.1. displays a situation where 2 individuals with different risk preferences, or (as Friedman and Savage have shown us) similar risk preferences but different amount of initial income, are faced with a decision to either take a certain amount of I or participate in gamble with expected utility of $U(\bar{I})$, and expected yielded income of $\bar{I}$.

The line between C and E represent possible utility resulting from the gamble with $U(I_1)$ representing the state of utility in case of lost and $U(I_2)$ corresponding to the state of utility in case when the gamble is won. The certain income with the same utility as the expected utility of the gamble is labelled as I^*. The concave/convex curve corresponds to the personal utility of the wealth function of the individuals participating. The graph a) displays concave shape of the utility function which consequently means that the person is risk-averse in this part of his utility function. The graph shows that the person who is risk-averse at this level of income will take a certain amount of I^* and maximizes his/her utility resulting from this gamble. The risk-seeking person will choose to gamble for expected value of $\bar{I}$ and maximizes his utility as well. The reader can notice a gap between I^* and $\bar{I}$. The difference between two in an absolute value is the amount of income that the person is willing to pay to insure himself/herself against the risk (risk-averse) or in the other case it represents amount of money a person is willing to pay to be allowed to gamble (risk-seeking). The hypothesis and subsequently developed Friedman-Savage utility function very well describes the changing tendency of people to respond to gambling incentives based on their income.

However, the hypothesis fails to explain continuous participation in lotteries even after accounting for risk premium. The same property that made the paper work in reality for a certain gambles and simultaneous purchase of the insurance fails to describe participation in a continuous gambling. If the reader takes a look at the figure 2.1 and pays attention to the "anchors" of the described function $U(I_2)$ or $U(I_1)$ respectively, he/she soon realizes that by performing one or more gambles a person will eventually arrive to a point where he/she no longer pleases to participate in gambling or to insure himself/herself, yet that contradicts what we experience in reality. There is significant amount of people who repeat the same type of gamble every week/month.

Friedman's former student Harry Markowitz suggested in his work – The utility of wealth (1952) that based on Friedman-Savage paper, the people in the highest level of income class should never participate in gambling or any sort of risk and subsequently large wealthy companies could divert certain portion of their money towards hedging against all kinds of risk, but that is not what is being observed. Funny enough, it is quite an opposite; wealthy companies invest into acquisitions of smaller ones rather than to their reserves and rich people playing poker in Las Vegas casinos against each other. Some might say that then the marginal utility of the wealth must be increasing in order to justify such behaviour, but Markowitz as one of the first economists suggested that the motivator for such behaviour must be a utility in the gambling itself. Something that Von Neumann and Morgenstern discarded in their independent axiom.

Upon the idea of the utility of the gambling did James Conlisk build his paper – The Utility of Gambling (1993). Therefore, the expected utility theorem even while accounting for the shifts in persons' preferences fails again. To solve this issue John Conlisk in his work proposes existence of utility from participating in the gambling. In this thesis this utility is referred to as lottery premium.

Furthermore, in 1980 Martin J. Bailey, Mancur Olson and Paul Wonnacott released their paper arguing that Friedman-Savage paper implies that marginal utility of income rises with income as it is a direct result of the continuous willingness to gamble as already stated above. Furthermore, they take a different look at the issue and argue that Friedman and Savage worked on their paper completely disregarding the literature on time preference and suggest that the two issues are related and once the time preferences or lending/borrowing and consumption distribution over time is taken into account the Friedman-Savage gambles are no longer the most feasible actions to take from the position of the individuals that have to decide how to distribute their income. They highlight 3 major points:

1) Stable levels of consumption are commonly observed when intertemporal utility is separable, therefore implying that the marginal utility of income decreases with rising income.
2) If we omit existence of point 1) and assume increasing marginal utility of income, gambles that Friedman and Savage mention in their paper do not maximize utility; incorporating the time preference reveals us that the individual would maximize his utility performing saving and dissaving actions and not participating in gambles unless there were severe imperfections in the capital market
3) Under circumstances when temporal separability of the utility function is not possible, the repeated participation in the gambling is not a rational way of maximizing utility from income under increasing marginal utility of income.

The paper does not disregard the gambling at all, but the authors point out that rational individual would only gamble in very specific conditions of the interest rate low enough or the expected utility of the yield of the gamble would have to at least equalize utility related to time preference (possible preference for uneven consumption in a certain time period) and the utility of the certain income that comes from the savings for that period.

This particular paper does not benefit the thesis in many ways but during its reading the author got an idea regarding the choices that he personally witnessed happening in the video game market. Certain people are hoarding the loot boxes and not opening them for a certain time period when they perform something so called “massive opening”. Even though, each box is independent event and therefore the result of opening 100 boxes over the course of 100 days or during 1 hour should not differ, people usually tend to buy larger amounts of the boxes together or save boxes for a specific period or spare their real/in-game money to purchase boxes in special action when new items are available or there is a sale resulting in people obtaining more boxes for the same amount of money. Therefore, author found parallel about the time preference on the gambling as a consumable good and decided to further investigate this phenomenon.

Chapter 3 Terminology, Definitions and Characteristics

The purpose of this chapter is for the author to help readers understand several specifics of the environment in which the research is performed. The first section of this chapter contains definitions of the basic terms as well as explanation of the terminology used throughout the thesis. The second section is dedicated to the description of the population upon which the data gathering process will be performed based on a research previously performed on a similar population. In the third section of this chapter, the author points out specifics of the market and all purchases that individuals are able to perform in it.

3.1 Terminology

Before the author presents readers with the definitions and the environment of the video game market it is necessary for the readers to understand meaning behind certain specific words used in this environment. The author decided use these words in the text of the thesis as well as in the questionnaire not only to maintain certain level of consistency, but also because they very well describe events/objects in one or two words that would be very problematic to replicate with regular vocabulary, would feel unnatural in the context of the thesis and wouldn't be able to capture the tiny nuances of the meaning. The author states that there has been no officially released thesaurus for the gaming environment that he would be aware of and therefore he relies on over 10 years of his experience in this field. The words that are essential to understand the thesis are following:

1) **Drop**
   Drop is a summary of all the virtual items that are obtainable in a specific way. That way being either defeating in game challenges (understand killing monsters e.g.) or is receivable from the loot box. The importance of this definition is that the drop is receivable by the individual neither already received nor has to be received at all. The origin of the word usage comes from the fact that since the dawn of videogames where the individual was represented by the character with an actual inventory (a bag into which you could put items that you found so that you could carry them and use them later) the foes you defeated along the way literally dropped a sack or the items themselves on the ground. Thus, players started calling it a drop. As already mentioned in the introduction, there is no law that would force the game producers to state the odds of different drops from loot boxes.
2) **Loot**
   Related to the drop, the loot represents the summary of items that player has actually received. Though the key item that the author is performing the research about is called loot box actually contains drop before it potentially becomes loot. The origins of the

word usage loot in this field are very similar to the drop one. Whenever player decided to let his character loot the items that dropped it became loot. The reason why the author highlights that there is a chance that the drop is shared between multiple individuals and can become loot of only one of them.

3) **Loot box**
   A virtual container (usually shape of a box) containing virtual items. The details are in the definitions section.
4) **Microtransaction**
   One of the monetization strategies implemented in the video game industry. The detailed explanation is in the definitions section.
5) **Subscription**
   There are two main types of subscription in the video game industry. The game only subscription is periodical (usually monthly) payment to maintain access to the game that uses this type of monetization. The service/platform subscription is a new trend on the video game market that is built on idea that majority of the players are bored with a specific game after playing it for few months/weeks and want to try new ones. This type of subscription gives you access to a variety of games for a limited amount of time, either total time spent playing the game or a real time limitation depending on type of the service that the individual is subscribed to and then rotates the game with the new subscription. One of these services that players can subscribe to is Xbox Game pass that incorporated model similar to Netflix with monthly payments. It gives people access to more than 100 games, whose list is being dynamically updated over time and the service has recently surpassed 10 million subscribers[13].
6) **Skin**
   In the video game industry, skin generally changes visual features of an item without influencing its other characteristics. Skins are one the drops that most games featuring loot boxes have in common. Simultaneously, they are also among the most sought after ones as they allow people to distinguish themselves from the rest of the people and customize their character. They come in different qualities with corresponding rarity. Nowadays people can apply different skins.
7) **Booster**
   Unlike its analog counterpart from the trading card games, the boosters in video games increase earnings of the game currency, experience that player is gathering in order to progress in the game or any other progression based commodity. Essentially, the individuals purchasing boosters are paying to increase their utility from every hour spent playing. In other words, people are paying so they do not have to play as much to reach the same in-game utility level.

[13] Source: https://www.zuora.com/2020/04/30/subscription-economy-news-week-of-04-27-20/

8) **Pay-to-win**
   The video games that are by players labelled as pay-to-win include microtransactions that cause unbalanced game experience for payers and non-payers. In other words if you compare to two equally skilled players from which one is payer and the other one is not in the environment of the pay-to-win game you would find out that the payer's performance will be always dominant.
9) **Meta**
   Meta stands for a dominant strategy in any type of action performed by the player in the game. More details are discussed in the Characteristics section of this chapter.

3.2 Definitions

3.2.1 Loot boxes

A loot box is the key term that is mentioned many times in this thesis. There are hundreds, possibly thousands of different types of loot boxes differencing in prices, shapes, content, rarity and a couple more aspects based on the game. However, despite all of the previously mentioned differences every loot box has five common properties. Those being:

1) It is a virtual container obtained via game interaction (purchased for real money, a game currency or simply by spending time in the game).
2) It contains one or more virtual items related to the game.
3) The content of the box is random, therefore unknown to the individual, drawn from the known pool of possible outcomes that varies in its size.
4) The items contained in the box do not follow uniform probability distribution, which unavoidably results in distinguishable rarity of the items.
5) Upon opening the box and revealing its content there is an eye-catching animation accompanied with sound effects

The general idea of a loot box is nothing new. Long before their digital forms were implemented into video games, there were physical versions of the boxes. It was first designed as a way of collecting physical items by the cigarette manufacturers in the USA in 1880s[14] that put the whole variety of collectible cards from the finite sets in their cigarette packages hoping that it will increase loyalty of the customers as they would want to finish the sets. All trading card games such as Pokémon: Trading card game, Magic: The gathering elevated the mere collecting into games of their own. These cards were no longer by-product of the cigarette package purchase but got the packages of their own called boosters. This was also the point where the rarity differentiation started taking place as there was set amount of common,

[14] Source: https://www.collectorsweekly.com/tobacciana/tobacco-cards

uncommon and rare cards per package. Certain cards from the early days are nowadays worth several thousand[15] US dollars, especially due to their rarity and condition.

3.2.2 Microtransactions

Before the author explains microtransactions, it is necessary to understand how monetization of the video games works nowadays. To be able to do that, one must first divide games into 3 general categories:

1) One-time purchase games
   Games that the consumer pays only to purchase and there are no subsequent obligatory payments to access the game.
2) Free to play games
   There are no obligatory payments at any point of time to access the game.
3) Games with periodical subscription
   These are games that require periodical (usually monthly) payments of the same sum to keep access to the game.

It is not uncommon that there are games that are 2 categories simultaneously, example being World of Warcraft from the Blizzard Entertainment, Inc., that is combination of 1) and 3). The importance of this basic categorization of the games is to recognize the main source of the income that the game generates. Distributors and developers of the 1) type games rely mostly on the amount of the copies purchased regardless of how long people play the game after the purchase. The monetization of the type 2) games consists completely out of voluntary microtransactions. The main stream of income for the games of the 3) type is the periodical payment that obviously scales with time and amount of people playing the game. Now the author can define the microtransactions.

As the term suggests, microtransactions are in-game purchases of virtual goods for real money. The “micro” part refers to the fact that these transactions are much smaller, ranging usually between 1-10USD/EUR, than usual price of a new AAA game, which is usually set to be equal to 59,99USD/EUR[16] for the PC platform for the one time purchase of the game. The microtransactions vary game to game as some are merely cosmetic and some give people real advantage in the game or increase their in-game progression speed. Among the most common ones belong skins, new playable characters, boosters and already mentioned loot boxes.

3.2.3 Market characteristics

The video game market can be considered quite young as it became mainstream entertainment in the 1970s with the game PONG and the first generation of video game consoles such as Atari. The industry is ever-changing as the trends are changing with every generation of players and obviously not all games are for all players. In the recent years the industry has seen a rise of the new trend wave, so called battle royal games. The concept is very simple and has been

[15] The infamous Black Lotus from Magic: The gathering:https://www.polygon.com/2019/3/5/18251623/magic-the-gathering-black-lotus-auction-price
[16] Source: https://gamingshift.com/60-dollar-aaa-video-games/

portrayed by movies and books. Certain amount of characters (controlled by other people) is deployed to an area that is shrinking in size as time progresses and the last person alive wins. The games are swift, allows for both low and high skill level players to satisfy their needs in the game and the loss has little consequences. Quite contrary to a long lasting, slow paced role-playing games where progress was the main motivator and ones actions had consequences. This type of games dominated in the middle-late 2000s. If by the change of mind-set and age of player base or technological advancement in the gaming industry, the way how people perceive gaming and what they expect from their gaming experience has dramatically changed.

The usual belief was that the diagram for the individuals spending their time playing video games was as follows: Pay to play and play to have fun therefore pay to have fun as it should be while spending their leisure time in a digital world. However, the perception of fun for players has changed over the course of years and is hard to tell if in a good way. Nowadays the diagram would look similarly to: Pay to play and play to win resulting in pay to win. People do not enjoy just simply spending time playing game, they want to win and succeed. Therefore over the course of last ten years the trend in the industry has shifted towards action multiplayer games with a competitive aspect – traceable rating representing performance of the individual. Due to rising competitiveness a new phenomenon has appeared – Meta gaming.

Meta has existed since the beginning of the gaming as every way of doing things, even in real life, has the most efficient way. One of the jobs of a good game designer is to ensure that all in-game mechanics are in a certain boundaries of normality, so called balance. It is impossible to achieve perfect balance and simultaneously maintain fun element that results from diversity in game. Thus, the Meta exists and awaits its discovery because surprisingly enough the games are so complex that their creators very rarely discover the dominant strategy before players themselves.

This leads to the mentioned phenomenon that is well described by the words of one the most famous game developers, Soren Johnson, who said: „Given the opportunity, players will optimize the fun out of a game.“[17] The Johnson argues that the player is having fun as long as he is convinced that he is spending his time efficiently. The fun lasts until he is told, shown or he himself discovers an exploit in the game design that leads to more efficient time spending. Unfortunately, such actions are usually very repetitive and not enjoyable for the individual performing it. But at that point of time he is unable to go back to his old methods and enjoy them the way he did before simply because his brain does not forget about a more efficient method therefore the individual is trapped by the utility maximisation and optimizes the fun of a game as Johnson predicted. The author himself has witnessed such behaviour and based on his experience can confirm that Johnson was right in general.

All of these properties of the gaming industry unavoidably lead to the fact that the fun factor of the engaging in the video has significantly changed its form. The existence of Meta was not

[17] Source: Game Developer Magazine, March 2011 issue, Column 17: Water Finds a Crack available at: https://www.designer-notes.com/?p=369

problematic at the early days as there was little interaction between players so the spread of the word and amount of people switching from their enjoyable game style to more efficient mind-numbing one was not so massive and sudden. In the multiplayer games is a certain feeling of necessity to follow Meta amplified by the fact that you face people who are already playing along the Meta and therefore are having an advantage over the individual. To remain competitive it is almost obligatory to follow this path, otherwise the individual could as well quit playing the game for good, but the fact is that people found out about the Meta with already large amount of time/effort possibly real money invested in the game. Then they perform the sunk cost fallacy as described by Arkes and Blumer (1985) and mostly escape this status quo by investing additional time or resources to remain competitive.

The author describes the thought process of the individuals spending their time in the video games in such detail because the video game manufacturers are aware of these motivations and other effects that influence players' decision making process and are monetizing upon it. One of the general characteristics of the games' economic design that have implemented microtransactions is that there are in most cases no open market operations within the game. Put into other words, once player charges real money on his/her in-game account, no matter which purchases he performs, the investment is unable to leave the account. The only possibility for a player to monetize his investment in real currency is to sell the account as a whole.

It is because of the difficulties that player has to face while trying to sell the account why most[18] people do not opt to perform this action. The first issue is that the accounts are bounded to the e-mails that are also used for the process of logging into the game and in case the player forgets his password he can recover it via his e-mail address. For the obvious reasons a possible buyer of the account would want to obtain the email account as well. The smart thing that the game developers do is that they bind the email address so it cannot be changed or can be changed through very annoying and long process of verifications. Based on author´s experience such process involves sharing your ID card and part of your bank transaction history with the game distributor.

They count on the fact that most people do not plan to sell the account or do not think about it at the moment of creation of an account. The second reason is that selling your account is against the terms of services that you have to sign before installing and loading the game for the first time. If a person happened to read the whole contract he would find out that the account is not his at any point of time and the operator (company) is giving him permission to access the account and such right to access can be revoked at any point of time. Game manufacturers nowadays track the position of the access in the world and if they find out a suspicious activity they will temporary, or permanently lock the account. The third reason is the risk of losing either account or the money.

[18] Based on the questionnaire findings.

As there is no official way how to sell an in-game account the individual must look for a buyer himself/herself on the game forums or third party sites on the internet. The direct transaction might be risky without a mediator as one of the sides has to fulfil their side of the deal first without guarantee (sending information of the in-game account and mail account or sending the money). To find a trustworthy mediator for two people who are strangers to one another is almost impossible task. The third party sites offer such service for a non-negligible fee ranging between 5% and 10% of the transaction, which removes most of the risk from the seller but reduces his profit from the transaction. Taking into account all of the risks and obstacles that a potential seller have to face, results in the author assuming that people paying for microtransactions cannot be considered to be investors. More on possible investment motivation, specifically loot boxes, is written in the methodology section of the thesis.

Chapter 4 DATA

This section focuses on the data collection, selection and preparation for the later analyses. Since the individuals are making purchases from the distributors/developers of the games, the author offered a mutual cooperation in exchange for a random sample of data from a certain period of the purchases for the microtransactions that would be anonymous. Unfortunately, no cooperation was established with any of the distributors (author suspects that due to GDPR or a fear of technological espionage, or general distrust towards the author). Therefore the author had to completely rely on the survey data that he gathered.

4.1 Gathering DATA

The author has decided to use Google questionnaire as it provides variety of customization of the questionnaire, including conditional parts of the survey and random factor that will be later utilized for the control group data gathering. The data were gathered in a period of January-October 2020. Fortunately, the author relied on digital form of the questionnaire and the data gathering process was not influenced by the COVID-19 outbreak and ordered social distancing. Due to concept of the questionnaire and authors connection in the video gaming community he was able to place first iterations on multiple forums targeting specifically people whose answers were relevant to his survey. However, the best environment to gather data was the live stream on the platform called Twitch. For brevity, Twitch is the streaming platform for any content creator, originally designed for video game players, but as time passed there were added categories for art creators, talk shows, etc. Anyway, Twitch, unlike Youtube, is focused on live streaming, meaning that the interaction with people is immediate.

Similarly to Youtube, Twitch is free-to-access, and shares a certain portion of the monetization strategy with it – video advertisements. However, the cooperation of the streamer and the platform is deeper than of those on Youtube, the streamer and the platform are sharing profits from all sources.

Those being:

1) Subscriptions – monthly payments of 5USD by a viewer to get access to unique features on the streamer's channel
2) Donations – from a certain amount of average viewers a streamer is able to establish donate connection to his PayPal or bank account and viewers can send the entertainer money to support him. This usually comes with ability to send a message with the amount donated that is displayed on the screen with a certain animation and is usually accompanied with a sound effect, which the streamer sets up. This is in general called Twitch alert.

The author used both of these methods to attract people to answer his survey. The usual routine that the author performed consisted of a finding a streamer playing one of the games containing loot boxes with a mediocre crowd of viewers (700-1500) people and the amount of money required to trigger the twitch alert upon donation (usually varied between 5-10USD/EUR). These three characteristics were chosen because people usually watch streamers playing the same games that they play themselves. Hence, as long as there were loot boxes in the game there was almost certainly somebody who purchases them. The amount of people chosen was a compromise based on several reasons. With the crowd of such size the streamer usually still reads all the donations even from the smaller amounts, also the amount is balance between a good chance of receiving several answers to the survey due to a significant population and the chat not moving too fast.

To clarify, people communicate with the performer via chat in texted messages, but since every streamer in this viewer count range has usually microphone as well as camera (the author checked all the streamers that satisfy the above specified category for camera and microphone at the moment of writing this part, the statement holds) he/she replies in a spoken way. When anyone sends a message the chat shifts not only for the streamer, but also for the other viewers, who are potential responders.

The process was supposed to work in the following way:

1) After all the observed criteria, that the author stated, are satisfied a donation was sent usually for the minimal amount to trigger the twitch alert. That notified not only the streamer which was the key part but also briefly the audience so they would pay attention for the next minute, which is crucial.
2) The donation was sent usually with the following text, there were some minor changes made during data gathering time, but no major changes. The text was: “Hello, nice stream you have there. I am in need of help from you and your followers. I am currently working on my bachelor thesis about loot boxes. Can you spam this link to a chat and ask your viewers to help me fill the survey? Thank you. “ (The link was at the end).
3) In almost all cases except 3 the link was sent into the chat (only the streamer can post URL links to the chat of his stream), and diverse population of people filled in the questionnaire.

Now the author will present how he created the questionnaire.

4.1.1 Creation of the questionnaire

The main goal was to design an ideal questionnaire that would provide the writer with enough data to perform reliable test and also short enough to not discourage people to complete it. In order to design such questionnaire the author must reconcile what he needs to obtain from the subsequent testing. The author needs to gather information on all variables that can possibly influence the lottery premium. Additionally he needs to capture the nature of the participants from the risk preference perspective. In order to achieve it, the author was looking for inspiration in literature on willingness to pay estimations (Christoph Breidert, Michael

Hahsler, Thomas Reutterer; 2006). That used survey as their data source as well as a paper on the methods of eliciting ones risk preferences (Gary Charness, Uri Gneezy, Alex Imas; 2013). Conlisk (1993) suggests that people are risk-neutral over small neighbourhoods at the starting point of the model. Therefore, even a small incentive can push otherwise risk-averse individual to the risk-seeking domain. Kahneman and Tversky (1979) also state that in accordance with the risk-averse hypothesis, people need substantively larger win to recover the lost utility. Furthermore, their model suggests significantly steeper utility function in the domain of losses. That implies that people have tendency to be risk-seeking in the domain of losses and risk-aversive in the domain of wins. The writer included another multiple price list to test if this theory works in the video-game environment as well. The values on the additional multiple price list represents values of losses to correspond with the mirror image result from the test that Kahneman and Tversky performed.

The questionnaire has two sections. First aims to gather general information such as age, education, disposable income, etc. These variables may or may not be significant in our model. The second part of the questionnaire is related to “testing”. From the several possibilities presented in the papers, the writer chose 6 separate sets of test questions. 2 related to profits, 2 to losses and 2 to loot boxes. Each of those contains 5 multiple choice questions, in which the respondent is presented with a risk-free option (its value is equal to the expected value of the median chance question from the given set) and chance involving option. Within each set the expected value of the chance involving choice is increasing/decreasing linearly. There is always one set of testing for lower values (5USD risk-free option and 10USD chance involving option) and one set for the higher values (10USD risk-free option and 100USD chance involving option). The sets for profits and losses are mirroring each other. As for the loot box related testing the author sets up environment for the respondents in which they can sell the item for its face value in exchange for real money. They are then presented with option to either buy a loot box that has chance to contain 10USD or respectively 100USD item or to purchase 5USD/10USD item from the in-game shop. To help capture the lottery premium effect the previous 4 test sets have no purchasing related to them. The latter 2 on the other hand are emphasizing the fact that person purchases the loot box. Also the author added two extra sets type-in type of questions related to loot boxes that allow respondents to manually type in the exact amount that they would be willing to pay for every single loot box option presented in multiple choice questions. The writer will elaborate onto this in methodology section.

4.2 Variables and Hypotheses

This section is dedicated to defining dependent and independent variables that the author will work with. The goal of the thesis is to determine which variables significantly influence amount spent on loot boxes per month and whether lottery premium is such variable or not. If it proves to be a significant variable the author will perform another regression with LP as a dependent variable to determine which variables significantly influence it.

4.2.1 Dependent variables

The author chose two dependent variables to be tested in this thesis. He performs OLS regression using *amount_spent* as the dependent variable in the first regression and *lottery_premium* in the second one.

Table 4.1: The list of dependent variables

Variable	Description
amount_spent	The amount spent monthly for a loot boxes
lottery_premium	The difference between willingness to pay for loot box purchase and pure money gain/loss alternative with the same odds.

4.2.2 Independent variables

Below are listed independent variables that are shortly described. DATA for those variables were collected through survey that the writer conducted. More details on those variables will be given in the hypothesis section.

Table 4.2: The list of independent variables

Variable type	Variable	Description
Continuous	*Age*	Years of age
	disposable_inc	Income of the respondent
	amount_spent	Amount of money monthly spent on loot boxes
	friends_open	Amount of friends who play with respondent at least once a week and of which majority buys loot boxes.
	box_item_value_root	The square root of the value of the most expensive item respondent received from loot box.
	lottery_premium	Quantified utility of loot box purchase
	hours_played	Amount of hours played/spent watching video games per week.
Dummy	box_opening	Dummy variable for watching loot box openings online
	odds_search	Dummy variable for research of probabilities before purchase
	Microtransactions	Dummy variable for other microtransactions purchases.
	casino_games	Dummy variable for casino games or other risk involving activities outside of the video games.
	account_sold	Dummy variable for ever selling an account by respondent.
	lootbox_purch_3m	Dummy variable for loot box purchase in last 3 months.
Categorical		
education_level		Categorical variable representing current/highest achieved educational by respondent.
	Elementary School	
	Grammar School	
	High School	
	University Undergraduate	
	University Graduate	
risk_lootbox		Categorical variable representing respondent behaviour towards loot box purchases from risk perspective.
	Risk seeking in high values.	Risk seeking only for the 10/100USD range.
	Risk averse	Risk averse over the whole range.
	Risk seeking	Risk seeking over the whole range.
	Risk seeking in low values.	Risk seeking only for the 5/10USD range.
lootbox_pur_freq		Categorical variable representing the frequency of lootbox purchases.
	Once a week	
	Once a month	
	Irregularly	
	Only during special events.	
payment_method		Categorical variable representing the favourite payment method.
	Online payment (debit card)	
	Paypal	
	Paysafecard	
	Mobile payment	

4.2.3 Hypotheses

In this section, the author presents hypotheses to be tested as well as reasoning why they should hold.

Hypothesis 1

The age positively influences amount spent.

The age at this specific environment is very important factor of decision making. Since people in young age are easier to be manipulated by their contemporary and by media. Author suspects small, but still significant effect on the amount spent as the respondents are getting older.

Hypothesis 2

The disposable income positively influences the amount spent.

Based on Friedman and Savage work people at a certain level of wealth should not participate in such gambles anymore. On the other hand, the writer suggests that people with larger income are willing to pay more to experience excitement from the game.

Hypothesis 3

The casino games positively influences the amount spent.

The author hypothesizes, that people spending their money in other risk involving activities are more likely to spend money/more money on the loot boxes.

Hypothesis 4

The average amount of time spent playing/watching video games per week positively influences the amount spent.

The person spending more time in game is more likely to encounter players wielding rare items obtainable only through loot box purchases. Therefore, referring to sunk cost fallacy, he is more likely to spend the more money the more he invests his time in the game.

Hypothesis 5

The odds search negatively influences the amount spent on loot boxes.

The loot boxes, much like lottery tickets, in the real world are not a game of fair chance. Hence, the author hypothesizes if the person takes time to investigate on the internet the true odds of the loot boxes, he will be less willing to pay/ pay as much for them as he would if he did not search at all.

Hypothesis 6

The amount of friends opening loot boxes positively influences the amount spent.

Based on the writer's personal experience people usually start playing/keep playing multiplayer games because of their friends. Thus, the friends shape the behaviour pattern of the respondents and the writer suspects it also has effect on the loot box purchase behaviour.

Hypothesis 7

The variable representing selling the account is not significantly influencing the amount spent.

The author believes that investment is not a significant motivator in loot box purchasing as he presents in definitions part of the thesis, therefore he suspects there is no significant relationship between amount spent and respondents selling accounts to monetize their "investment".

Hypothesis 8

The grammar school, high school and elementary school have significant positive influence on the lottery premium.

The author suspects that the respondents that are not yet university students/never studied at a university will be more likely to have higher lottery premium.

Hypothesis 9

The respondents with risk seeking behaviour, in loot box purchase environment, positively influence the lottery premium.

The author justifies this hypothesis partially by the fact how the lottery premium is calculated in this thesis and also by the fact that respondent is marked as risk seeking if he chooses chance involving choice over the certain one at the threshold option where expected value of both equals. That is in line with the general idea of expected utility theory. The risk seeking person chooses the option with higher variance over the certain option; given the both have the same expected value.

Hypothesis 10

The microtransactions variable has a positive influence on the amount spent.

The respondents that purchase other forms of microtransactions are more likely to also purchase loot boxes, since they are already spending money in game.

Hypothesis 11

The lottery premium has significant and positive influence on amount spent.

The author hypothesizes that lottery premium is able to justify the decision making process of the individuals purchasing loot boxes.

Hypothesis 12

The root of box item value has a significant positive influence on amount spent.

The author hypothesizes that previous win has an impact on future decisions, similarly to previous loss. In other words, if the respondent got something valuable in the past it affects his amount spent now or in the future. The reason why the author uses the square root is explained in the results section.

Hypothesis 13

The loot box purchase in last 3 months does not have a significant influence on the amount spent.

Based on the author's experience, people usually tend to buy more boxes at the same time but less often, for example once or twice a year during special sales or actions. Therefore, the answer is irrelevant, if such action/sale had not taken place in the 3 months period before the respondent answered the questionnaire.

Hypothesis 14

Watching a loot box opening stream or video has a positive effect on lottery premium.

The writer hypothesizes that watching loot box opening does not directly influence the amount spent, because the other factors are suspected to have stronger influence. However, he believes that it could affect the size of the premium. The reasoning behind this suspicion is that watching other people open loot boxes and win might influence the risk attitude towards their purchase.

4.3 Descriptive statistics

Table 4.3 shows the descriptive statistics of the chosen dependent variables.

Table 4.3: Descriptive statistics of the dependent variables

Statistics	Mean	St. Dev.	Min	Max
amount_spent	12,108	7,451	0	35
lottery_premium	1,367	16,067	-29,5	33

N=166

Unfortunately, the sample of collected data is smaller than the writer planned because of the fact that the questionnaire had 3 iterations and the data from the previous iterations would not benefit the current models. Also, the author's budget for the data gathering was fully expended.

The distribution of the amount spent, as presented in figure 4.1, shows that the amount most people spent monthly on loot boxes is within 0-15USD range. Which is in a line with Conlisk model where people are risk neutral at the small neighbourhood at the small values and only small incentive is required to make them do risk seeking decision. The author suggests that people spending money within this neighbourhood can be heavily influenced by the factors presented in the models later.

The distribution of the lottery premium variable can be seen in figure 4.2, it shows that the majority of the values is on the positive side of the distribution, which is corresponding with the author's theory of the lottery premium influence, as the data were collected only from the people previously purchasing the loot boxes.

The following table 4.4 shows descriptive statistics for all the relevant independent variables used in the models. Majority of the variables are binary in their nature or transformed into the binary variable from the categorical ones for the purpose of this statistics. The mean for such variables stands for their proportional representation in the collected sample.

Figure 4.1: Histogram of amount spent

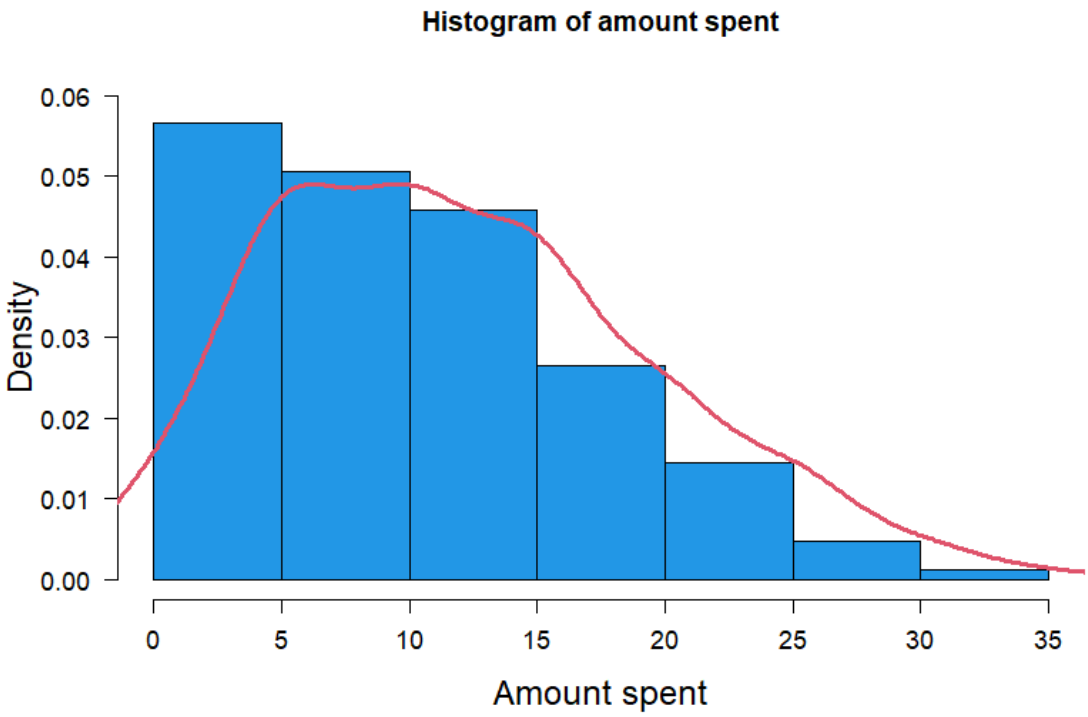


Figure 4.2: Histogram of lottery premium

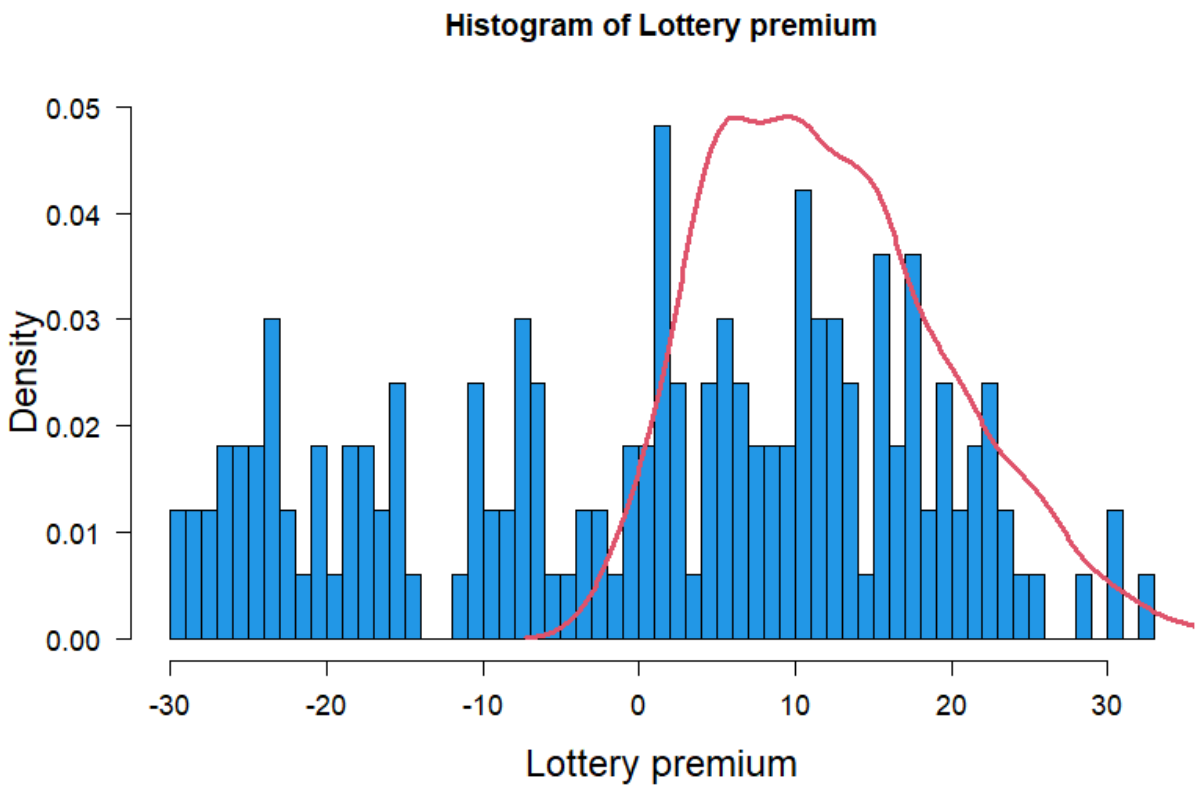


Table 4.4: Descriptive statistics of independent variables

Statistics	Mean	St. Dev.	Min	Max
Continuous				
Age	26,93	5,06	13	41
disposable_income	527,14	437,65	60	2200
Box_item_value_root	6,57	3,69	1	20
Friends_open	2,20	1,87	0	5
Dummy				
odds_search	0,34	0,47	0	1
Box_opening	0,38	0,49	0	1
Educ_elementary	0,07	0,26	0	1
Educ_high_school	0,19	0,39	0	1
Educ_grammar	0,05	0,21	0	1
Educ_uni_under	0,25	0,43	0	1
Casino_games	0,56	0,50	0	1
Box_pur_freq_month	0,13	0,34	0	1
Box_pur_averse	0,21	0,41	0	1
Box_pur_seeking	0,53	0,50	0	1
Box_pur_seeking_low	0,22	0,42	0	1

Chapter 5 Methodology

This chapter is dedicated to the description of the models and the dependent and independent variables that are used for estimation. The author uses the OLS model in order to determine significant variables and how much they influence the amount spent. The theory of modelling of this section is based on Wooldridge (2015).

5.1 Computing the lottery premium

As the lottery premium is essentially a quantified difference between the same gamble without any associated joy from the loot box purchase and opening process and without a deeper knowledge whether respondents treat loot boxes as potential loss or potential profit he decided to take expected value from the first 4 sets of questions mentioned in the DATA section sum them up and divide them by 2 to create a mean between losses and profits. Then the author takes sum from the last segment of testing, where respondents filled in their willingness to pay for every specific loot box presented and subtracts the previously found profits/losses mean from it. The resulting number represents the lottery premium.

5.2 OLS model

In this section, the author presents the assumptions of the OLS model and justifies their upholding in the presented models. The author presents several models in the following form:

$$y_i = \beta_0 + X\beta + u_i,$$

where y_i represents the dependent variable with respect to the *i*-th respondent, amount_spent in our first OLS model, β_0 represents intercept of the model, X represents the vector of independent variables and u_i is the error term which author assumes to be identically independently and normally distributed.

For the model to be valid, the author must now test the Multiple Linear Regression assumptions (MLR.1 to MLR.5) that must be satisfied. The writer stated those assumptions in the Appendix A. The author will use visual examine of the relevant plots to determine whether the assumptions are satisfied or not.

As for the linearity assumptions, the author investigates Residuals vs Fitted values plot and finds the line within acceptable boundaries satisfying linearity, also all the parameters presented in the model are linear. By adding as many variables as possible the author is controlling for no correlation of error term with independent variables. As for the random sampling assumption, the data are collected throughout random communities of people with similar hobby. The normal distribution of residuals is visually examined from Q-Q plot attached

in the Appendix B and the author finds residuals normally distributed. As for homoscedasticity assumption, the author justifies the “patterns” in the Residual vs Fitted values and Scale-Location plots by a limited amount of values that dependent variable takes. The explanation is based on the paper written by Shayle R. Searle (1987).

Therefore the writer hypothesizes that the parallel lines present are not related to the heteroscedasticity. The author performs Breusch-Pagan test and does not refuse homoscedasticity, the test result is attached in Apendix. The writer also investigates the Residuals vs Leverage plot for any outliers and influential cases that could possibly influence the model. Finally, to check for collinearity between independent variables we use Variance Inflation Factor method. Therefore, under the assumptions (MLR.1 to MLR.5) the OLS estimator is the best linear unbiased estimator.

Chapter 6 Results

This chapter presents the results of regressions and confirms/refuses the hypotheses that were stated earlier. The primary software that helps us analyse and estimate data was the R software. First the author presents the regression results. We performed multiple OLS regressions to find the best fitting ones for amount_spent and consequently lottery_premium. The author presents the amount_spent OLS model first.

6.1 Regression results – The OLS model (amount_spent)

The author performs the first regression with amount_spent as dependent variable. To fight the patterns mentioned earlier, the author tried different transformations of the variable but as expected not a single transformation helped. Furthermore the author performed transformation of one the independent variables (box_item_value) into its square root to help fight collinearity and also the data distribution. Based on the nature of the loot boxes, there was extensive amount of respondents reporting very low numbers.[19]

To find the best fitting model the author used adjusted R^2 as a metric while simultaneously keeping the significant variables not suffering from multicollinearity in the model. As stated before the author runs B-P test to check for homoscedasticity assumption and VIF test to account for collinearity assumption using James et al. (2013) as reference. Their results can be found in Appendix B. The final model representing the amount spent per month looks as follows:

$$amount\_spent = \beta_1 age + \beta_2 casino\_games + \beta_3 odds\_search + \beta_4 friends\_open + \beta_5 box\_item\_value\_root + \beta_6 disposable\_inc + \beta_7 Lottery\_premium + \varepsilon .$$

[19] Based on the nature of the loot boxes there are more people with small wins than with, more impactful, bigger wins. Therefore shifting the whole scale towards the smaller values helps with more even distribution of values without impacting the significancy.

The following table represents the results of the regression:

Table 6.1: OLS regression (amount spent)

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	-2.310	1.833	-1.260	0.210
age	0.209	0.074	2.829	0.005 **
casino_gamesYes	2.909	0.712	4.084	7.02e-05 ***
odds_searchYes	-2.283	0.624	-3.658	0.000 ***
Friends_open	0.426	0.166	2.570	0.011 *
box_item_value_root	0.642	0.084	7.625	2.13e-12 ***
disposable_inc	0.005	0.001	5.530	1.30e-07 ***
Lottery_premium	0.186	0.022	8.617	6.72e-15 ***

Residual standard error: 3.757 on 158 degrees of freedom

Multiple R-squared: 0.7565, Adjusted R-squared: 0.7457

F-statistic: 70.13 on 7 and 158 DF, p-value: < 2.2e-16

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

6.2 OLS model (lottery premium)

As for the lottery premium OLS model, all the same procedures are performed. That includes (MLR.1 to MLR.5). The linearity is once again visually examined from Residuals vs Fitted values plot as well as normal distribution of residuals being investigated from Q-Q plot. There are no outliers or influential cases and the author finds no pattern in the Scale-Location plot that would possibly could mean heteroscedasticity. The author runs B-P test to confirm this observation and once again does not refuse homoscedasticity. The writer performs VIF test to check that there is no linear relationship between the independent variables, no such relationship is found and therefore this model satisfies necessary requirements to be considered BLUE.

6.3 Regression results – the OLS model (lottery_premium)

The author performs the second regression with lottery_premium as dependent variable. The independent variables used in this particular model are the ones that are not present in the first regression maintain assumption of no collinearity intact. The author uses the same principle as with the first model, running several regressions with different combinations of variables looking for the best fitting model, while maintaining the significant variables in the model and complying with the MLR assumptions.

As in the first case scenario, the writer performs B-P test on top of the visual examination of the plots to search for signs of heteroscedasticity and VIF test to determine whether the no collinearity assumption is violated. Both the tests and the plots can be found in the Appendix B as well as certain regressions to confirm hypothesis stated earlier. The final model representing the lottery premium of an individual looks as follows:

$$lottery\_premium = \beta_1 risk\_lootbox + \beta_2 education\_level + \beta_3 boxopening + \beta_4 lootbox\_pur\_freq\_month + \varepsilon$$

For the risk_lootbox the author sets reference level as “Risk seeking in high values” and for the education_level as “University Graduate”.

Table 6.2: OLS regression results (lottery_premium)

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	1.328	3.260	0.407	0.684
risk_lootboxRisk Averse	-14.186	3.581	-3.961	0.000 ***
risk_lootboxRisk seeking	7.022	3.342	2.101	0.037 *
risk_lootboxRisk seeking in low	-8.759	3.478	-2.518	0.013 *
education_level Elementary	11.423	3.253	3.512	0.001***
education_levelGrammar	-7.261	3.311	2.193	0.030 *
education_levelHigh School	6.573	2.123	3.096	0.002 **
education_levelUniversity under.	-7.313	1.838	-3.978	0.000 ***
boxopeningYes	5.482	1.709	3.207	0.002 **
lootbox_pur_freq_month	-6.059	2.274	2.664	0.009 **

Residual standard error: 8.381 on 156 degrees of freedom

Multiple R-squared: 0.743, Adjusted R-squared: 0.7279

F-statistic: 50.04 on 9 and 156 DF, p-value: < 2.2e-16

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

The author selects only "Once a month" option from the categorical variable to represent the frequency of loot box purchases, because it was the only significant part of the variable and its result is meaningful. Thus, the writer transforms it into the dummy variable, where reference level is set to "No" represented by "0" value.

6.4 Results – Hypotheses

The interpretation of the results from OLS regression is pretty straightforward. The results for both OLS regressions are found in the Estimate columns of table 6.1 and 6.2. Using the estimate values and the p-values, the author can investigate the hypotheses stated earlier. We set our significant level threshold at 5%, or 0,05 p-value.

Hypothesis 1

The age positively influences the amount spent.

The results from the regression support the stated hypothesis. As it is a level-level type of relationship between the dependent and independent variable, the effect assuming ceteris paribus of 1 year of age increase results into extra 0,21USD spent per month. The variable is statistically significant at 1% significant level.

Hypothesis 2

The disposable income positively influences the amount spent.

The results support this hypothesis. The author chose disposable income on purpose, because the author suspected that people would not be as compliant when asked for the general income. Also it is tied to a practical reason. There is no need to adjust for other effects related to the income, such as average salary in the respondents' country, price level, etc. The variable is statistically significant at 0,1% level. It represents level-level relationship and the effect of adding 1USD to disposable income is an increase of 0,005USD on monthly spending, ceteris paribus.

Hypothesis 3

The participation in a risk financial risk involving activity positively influences the amount spent.

The results support the author´s hypothesis. Participation in other risk involving activities outside of the video game environment has a positive and statistically significant effect on the amount spent. The result makes sense as certain people perceive loot boxes as a virtual

roulette. The statistical significance of the variable is at 0,1% level. The effect of the variable is quantified as 2,9USD of monthly amount spent on loot boxes if the person participates in risk involving activities, assuming ceteris paribus.

Hypothesis 4

The average amount of time spent playing/watching video games per week positively influences the amount spent.

Surprisingly enough, the effect of the hours spent playing/watching streams or videos is not statistically significant at any level. The author justifies this by presenting two effects that are related to time spent playing the game. There are people that spend money on loot boxes because they have invested a lot of time and money into the game and then there are people that cannot dedicate that amount of time to the game, but still want to remain competitive so they spend the extra money on the loot boxes to literally keep up with the rest. This is probably the reason why the regression does not show any significance. Therefore, this hypothesis is refused.

Hypothesis 5

The odds search negatively influences the amount spent on loot boxes.

The results of the regression show that the hypothesis holds. The author further hypothesizes that people that do not search for the odds are actually overvaluing their chance to win, but this idea will need further economic research. The variable is statistically significant at 0,1% level and it has negative effect which, ceteris paribus, decreases the amount spent monthly by 2,28USD if the person is searching for the loot box odds before purchasing.

Hypothesis 6

The friends opening loot boxes positively influences the amount spent.

This hypothesis is supported by the results. The statistical significance is at 5% level, which is our threshold. The variable has a positive effect that increases the amount spent per month by 0,43USD, ceteris paribus, if respondent gains a friend that buys loot boxes.

Hypothesis 7

The variable representing selling the account is not significantly influencing the amount spent.

This hypothesis holds, as predicted by the author. Therefore, the author's statement about investment not being a significant motivator behind the loot box purchases holds.

Hypothesis 8

The grammar school, high school and elementary school have significant positive influence on the lottery premium.

This hypothesis is only partially true. The variables increasing the lottery premium are the high school and elementary school. The effect of a grammar school variable is negative. The author hypothesizes that since grammar school represents the best education level from this group, it is beyond the threshold, where more rational approach towards the purchasing loot boxes. The effects are as follows: the high school variable is statistically significant at 1% level and it increases the lottery premium by 6,57 points if the person is studying/his highest achieved education is high school. It is similar for the remaining variables. The elementary school is significant at 0,1% level and its effect is 11,42 points increase in the lottery premium. The variable representing the grammar school has a negative effect that is significant at 5% level and has size of effect equal to decrease of 7,26 point of lottery premium. Besides these, the results also show that variable representing university undergraduate is significant at 0,1% level and its effect has negative impact on lottery premium. To be specific the effect is decrease of 7,31 points of lottery premium, if the variable is true.

Hypothesis 9

The respondents with risk seeking behaviour, in loot box purchase environment, positively influence the lottery premium.

The results of the OLS regression show us that this hypothesis is true. The risk seeking behaviour detected by the questionnaire is a significant variable in determining the lottery premium value. Its significance is at 5% level and its impact on the lottery premium is positive and has size of 7 points increase, ceteris paribus.

Hypothesis 10

The microtransactions variable has a positive influence on the amount spent.

The regression does not show a statistical significance at any level between amount spent and microtransactions purchases. Therefore the writer refuses this hypothesis. The author hypothesizes that the reason for insignificance of this variable is that certain people are using microtransactions as a substitute to the loot box purchases and a group of different people as a complement.

Hypothesis 11

The lottery premium has significant and positive influence on amount spent.

The first OLS regression confirms this hypothesis. This fact justifies the validity of the second OLS regression and the reason for its existence. The lottery premium is statistically significant at

0,1% level and has a positive effect on amount spent. The effects magnitude, ceteris paribus, is equal to 0,19USD increase if the lottery premium increases by 1 point.

Hypothesis 12

The root of box item value has a significant positive influence on amount spent.

The results show that the variable is significant at 0,1% level. The transformation was necessary to make the distribution of the values "more normal like". Furthermore, the effect of the variable is positive and equal to 0,64USD amount spent per month increase for a one point increase in the square root value.

Hypothesis 13

The loot box purchase in last 3 months does not have a significant influence on the amount spent.

The regression results confirm the hypothesis. The variable has no statistically significant effect on the amount spent for the reasons that the writer presented in the DATA section.

Hypothesis 14

Watching a loot box opening stream or video has a positive effect on lottery premium.

The results from the regression confirm this hypothesis. The effect is statistically significant at 1% level and its magnitude is estimated to be 5,48USD when the respondent indeed watches such videos. This indirectly means that watching loot box opening videos shifts respondents' risk attitude for loot box purchases more towards risk seeking side.

Chapter 7 Conclusion

The purpose of the thesis was to determine and analyse the significant factors influencing the amount spent, in particular to experimentally extract lottery premium from the answers of the respondents and test its significance as an independent variable towards the average amount spent per month for the loot boxes. The author has shown that there is a strong relationship. Therefore, the writer showed that the Conlisk's (1993) theory about the utility from the gambling itself, being one of the determinants why people participate in lotteries, holds. The author then ran the second regression to determine factors influencing the size of lottery premium. The thesis was motivated by the author's interest of analysing the video game industry from the perspective of an economist and the affairs that were related to loot boxes in prior years.

The author decided to use OLS regression for both dependent variables as it is a common procedure for similar type of research. The results obtained were mostly in line with what the author hypothesized and presented in the methodology section. We have determined that besides lottery premium, the next strongly influential factors determining the amount spent are: the value of the most expensive item obtained from the loot box, the fact if the person searches for the odds of the loot boxes before purchasing it and if the person participates in the risk involving activities outside of the video game environment. All the variables introduced in the models had an expected direction of the effect, even though some of them, such as hours played, did not show a statistical significance. As a result the author accepted most of the presented hypotheses.

Nevertheless, some of the hypotheses could be elaborated on further, especially on the odds search before the purchase. As the author stated earlier it is not a common practise to display odds of the loot boxes as a video game manufacturer. As a matter of fact there is no law forcing the companies to do so. If the odds are unknown and people still purchase the loot boxes regardless, how do they determine the expected value? The author hypothesizes that these people probably disregard the importance of the expected value or they base their decision making process on the artificial expected value of their own which is based on the watching other people open the boxes/hear from their friends what they received. It presents a problem; as such process of obtaining expected value can be very biased for multiple reasons. One of them being that most videos of box openings are the videos of people winning something, not many people publish video of winning nothing. It is much like an interview with a lottery winner; the person sees the winner but not the other people that have won nothing. Nevertheless, as the author previously mentions, this issue asks for further research in the field.

For the future research, the author recommends to select a sample of people and track their decision making process over the course of an extended time period, to see how their behaviour patterns change over time. The other possibility would be to conduct a treatment comparison research for possible government interventions into the market such as additional taxation on the loot boxes.

Bibliography

Wooldridge, J. M. (2015), *Introductory econometrics: A modern approach*, Nelson Education.

Von Neumann, J. and Morgenstern, O. (2007) Theory of Games and Economic Behavior : 60th Anniversary Commemorative Edition. Princeton: Princeton University Press (Princeton Classic Editions) pp. (28, 629,632)

Prokop, Darren. "Von Neumann–Morgenstern utility function". *Encyclopedia Britannica*, 15 Mar. 2016, https://www.britannica.com/topic/von-Neumann-Morgenstern-utility-function.

Blavatskyy, Pavlo (April 2005). "Back to the St. Petersburg Paradox?" (PDF). *Management Science*. 51 (4): 677–678

Arrow, Kenneth J. (February 1974). "The use of unbounded utility functions in expected-utility maximization: Response" (PDF). *Quarterly Journal of Economics*. 88 (1): 136–138.

Markowitz, Harry. “The Utility of Wealth.” *Journal of Political Economy*, vol. 60, no. 2, 1952, pp. 151–158.

CONLISK, JOHN. “The Utility of Gambling.” *Journal of Risk and Uncertainty*, vol. 6, no. 3, 1993, pp. 255–275.

Olson, Mancur, and Martin J. Bailey. “Positive Time Preference.” *Journal of Political Economy*, vol. 89, no. 1, 1981, pp. 1–25.

Bailey, Martin J & Olson, Mancur & Wonnacott, Paul, 1980. "The Marginal Utility of Income Does not Increase: Borrowing, Lending, and Friedman-Savage Gambles," American Economic Review, American Economic Association, vol. 70(3), pages 372-379, June.

Milton Friedman , L. J. Savage “The Utility Analysis of Choices Involving Risk“, Journal of Political Economy, Vol. 56, No. 4 (Aug., 1948), pp. 279-304

Markowitz, Harry. “The Utility of Wealth.” *Journal of Political Economy*, vol. 60, no. 2, 1952, pp. 151–158.

DIECIDUE, ENRICO, et al. “The Utility of Gambling Reconsidered.” *Journal of Risk and Uncertainty*, vol. 29, no. 3, 2004, pp. 241–259.

Arkes, Hal & Blumer, Catherine. (1985). The psychology of sunk cost. Organizational Behavior and Human Decision Processes. 35. 124-140. 10.1016/0749-5978(85)90049-4.

Breidert, Christoph & Hahsler, Michael & Reutterer, Thomas. (2015). A Review of Methods for Measuring Willingness-to-Pay. Innovative Marketing. 1.

Charness, Gary & Gneezy, Uri & Imas, Alex. (2013). Experimental Methods: Eliciting Risk Preferences. Journal of Economic Behavior & Organization. 87. 43–51.

Golec, Joseph and Maurry Tamarkin. (1998). “Bettors Love Skewness, Not Risk, at the Horse Track,” *Journal of Political Economy* 106, 205–225.

Fishburn, Peter C. (1970). *Utility Theory for Decision Making*. New York: Wiley.

Fishburn, Peter C. (1980). “A Simple Model for the Utility of Gambling,” *Psychometrika* 45, 435–448.

Schmidt, Ulrich. (1998). “A Measurement of the Certainty Effect,” *Journal of Mathematical Psychology* 42, 32–47.

Tversky, Amos. (2000). Prospect Theory: An Analysis of Decision under Risk. 10.1017/CBO9780511803475.003.

Chateauneuf, Alain & Cohen, Michèle. (1994). Risk Seeking with Diminishing Marginal Utility in a Non-expected Utility Model. Journal of Risk and Uncertainty. 9. 77-91. 10.1007/BF01073404.

Chiles, Todd & Mcmackin, John. (1996). Integrating Variable Risk Preferences, Trust, and Transaction Cost Economics. The Academy of Management Review. 21. 73. 10.2307/258630.

Shayle R. Searle (1988) Parallel Lines in Residual Plots, The American Statistician, 42:3, 211, DOI: 10.1080/00031305.1988.10475569

Thaler, Richard & Johnson, Eric. (1990). Gambling With the House Money and Trying to Break Even: The Effects of Prior Outcomes on Risky Choice. Management Science. 36. 643-660. 10.1287/mnsc.36.6.643.

Ng, Yew-Kwang. (1965). Why do People Buy Lottery Tickets? Choices Involving Risk and the Indivisibility of Expenditure. Journal of Political Economy. 73. 530-530. 10.1086/259077.

Gul, Faruk. “A Theory of Disappointment Aversion.” *Econometrica*, vol. 59, no. 3, 1991, pp. 667–686.

O'Donoghue, Ted, and Jason Somerville. 2018. "Modeling Risk Aversion in Economics." *Journal of Economic Perspectives*, 32 (2): 91-114

C. Faustino, Horácio & João Kaiseler, Maria & Marques, Rafael. (2009). Why Do People Buy Lottery Products?.

John Fox & Georges Monette (1992) Generalized Collinearity Diagnostics, Journal of the American Statistical Association, 87:417, 178-183

Appendix A: Plots and Assumptions

The Gauss-Markov Assumptions for multilinear regression

Under the asumptions MLR.1 - MLR.5, the estimator is BLUE.

MLR.1. Linear in Parameters
The equation of the model is always linear in parameters.

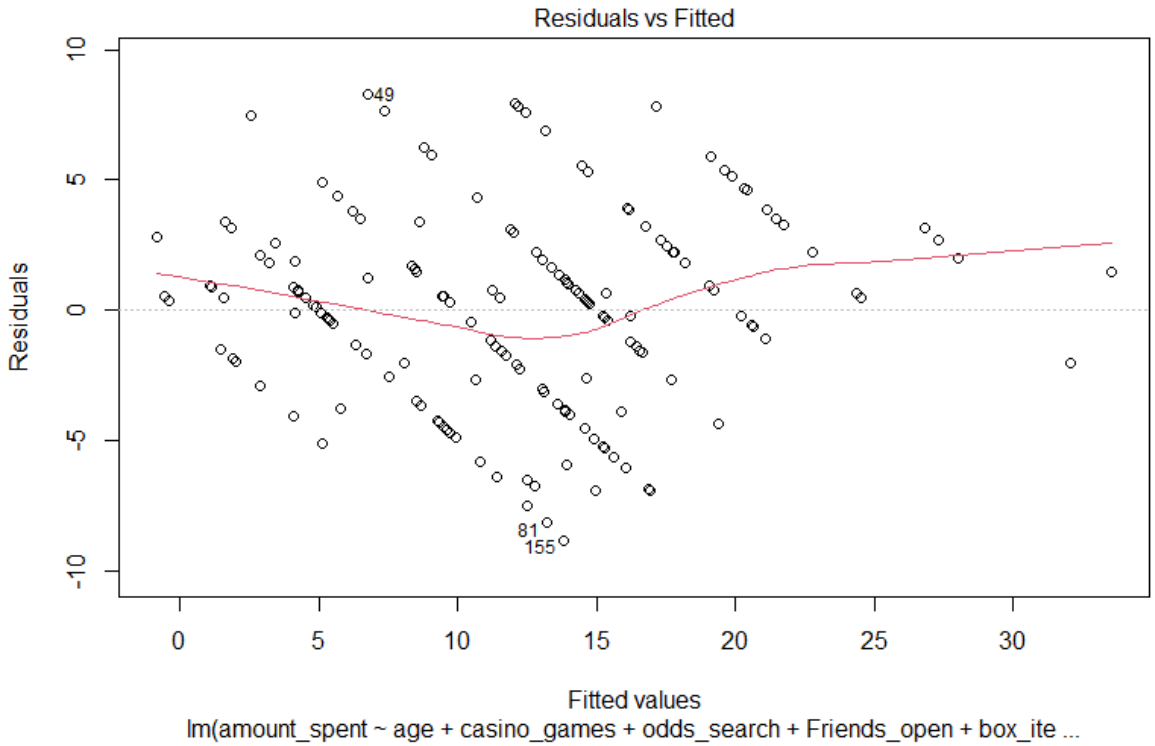
MLR.2. Random Sampling
We have a random sample of *n* observations following the population model from Assumption MLR.1.

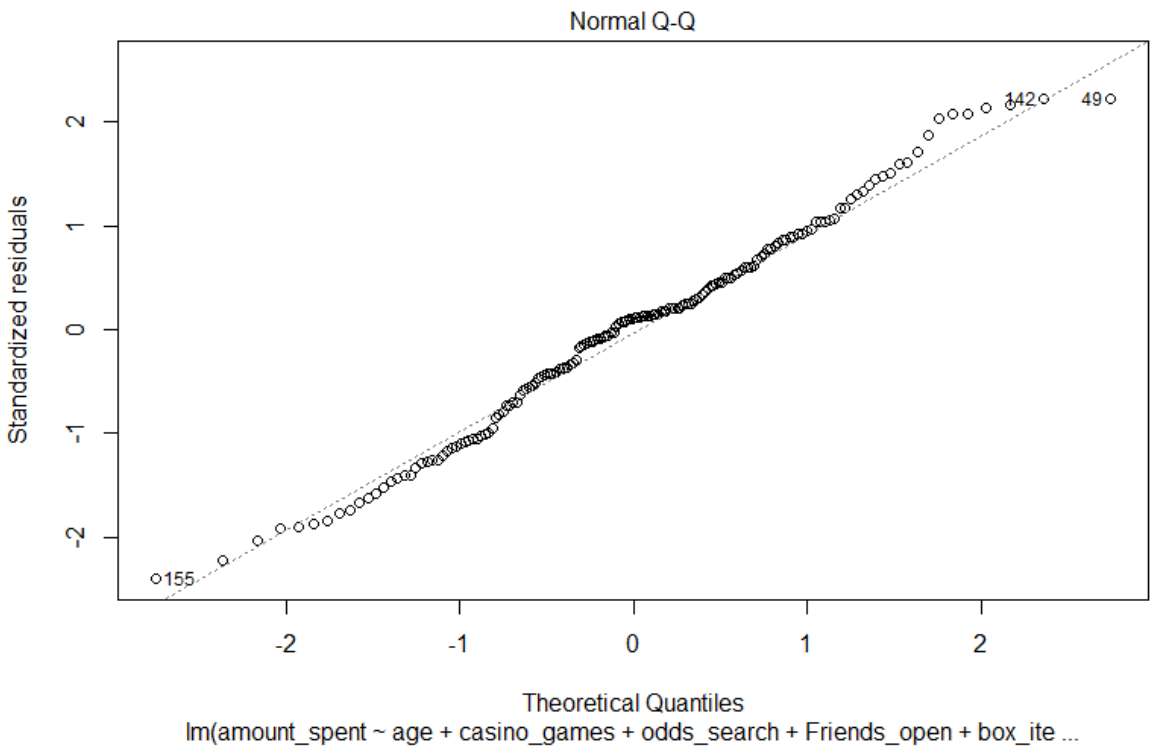
MLR.3. No Perfect Collinearity
None of the independent variables is constant and there are no exact linear relationships among them.

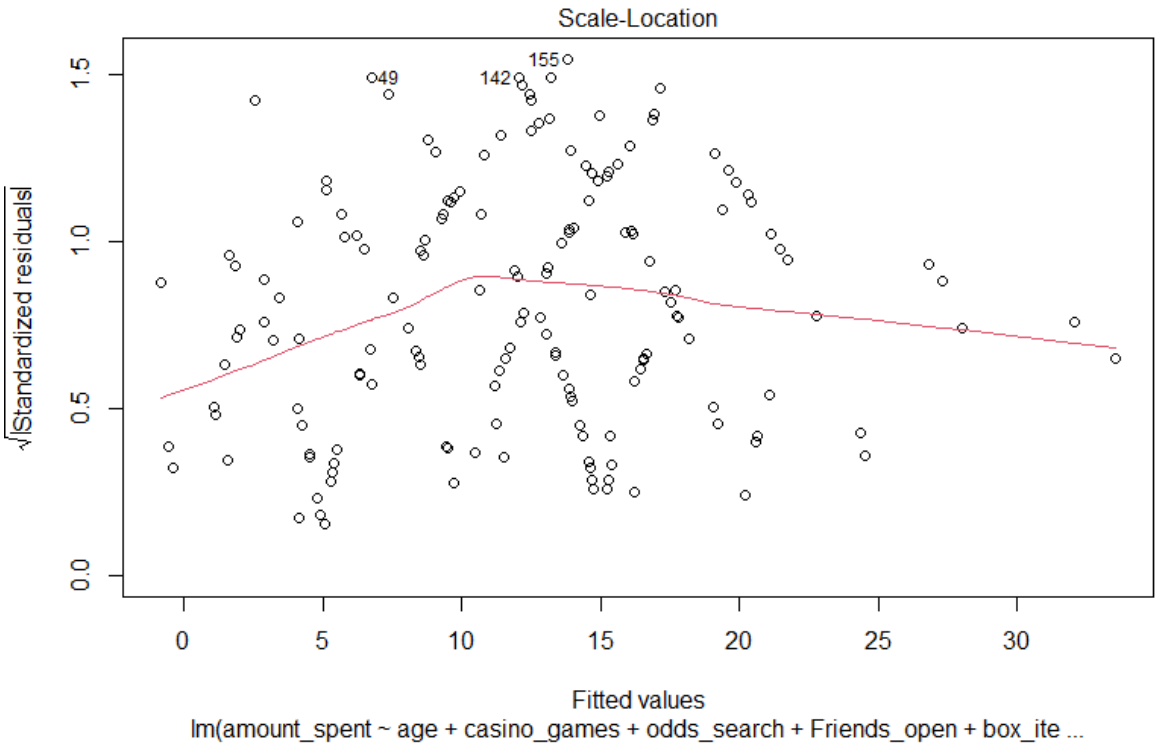
MLR.4. Zero Conditional Mean
The error term *u* has an expected value of zero given any independent variables.

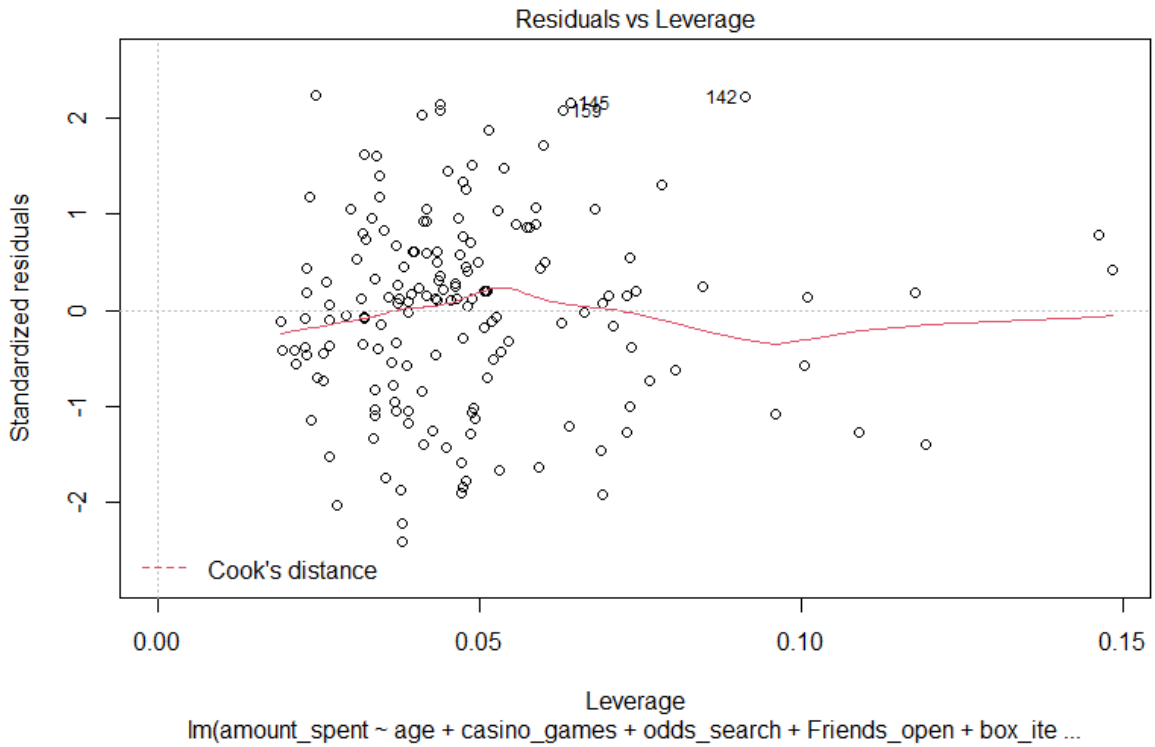
MLR.5. Homoskedasticity
The error term *u* has the same variance givenany values of the explanatory variables.

Plots generated by R from linear regression for amount_spent

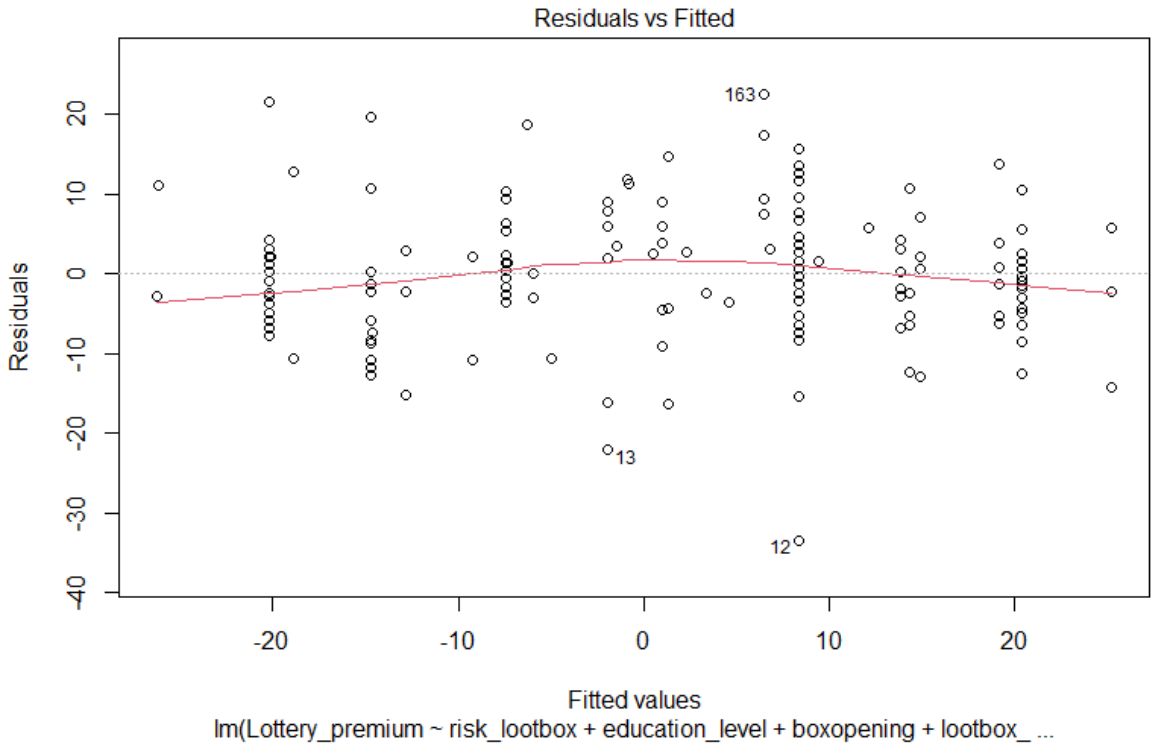


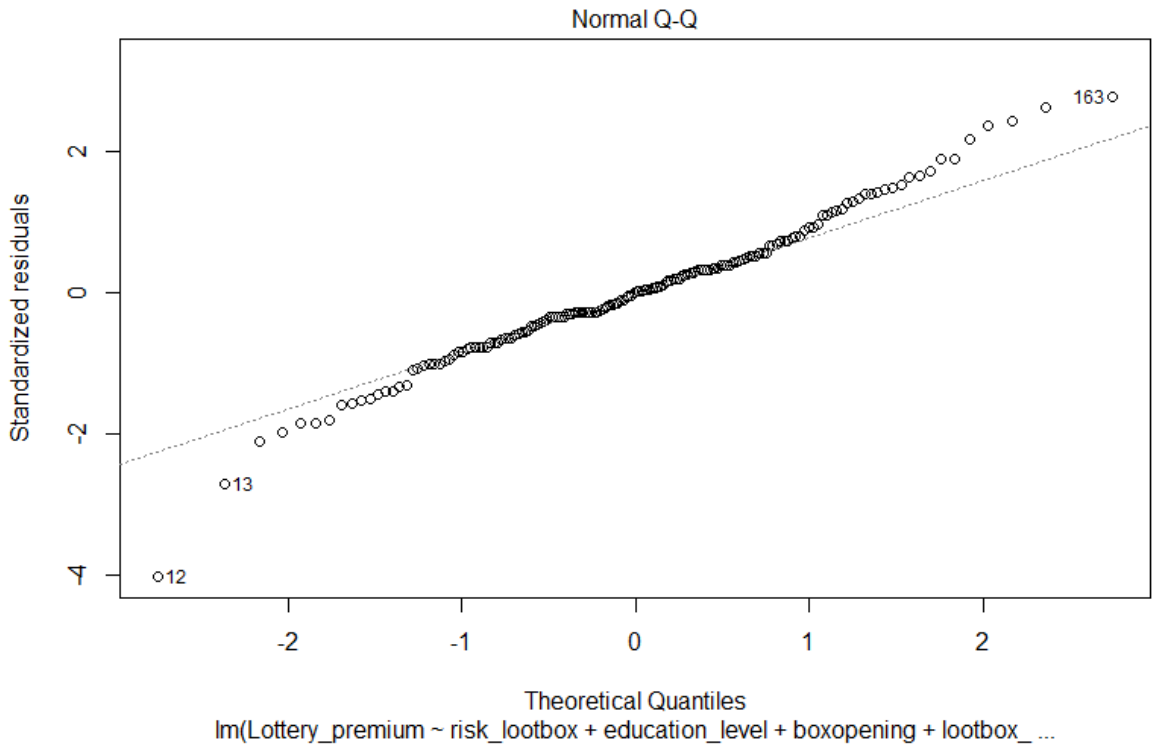


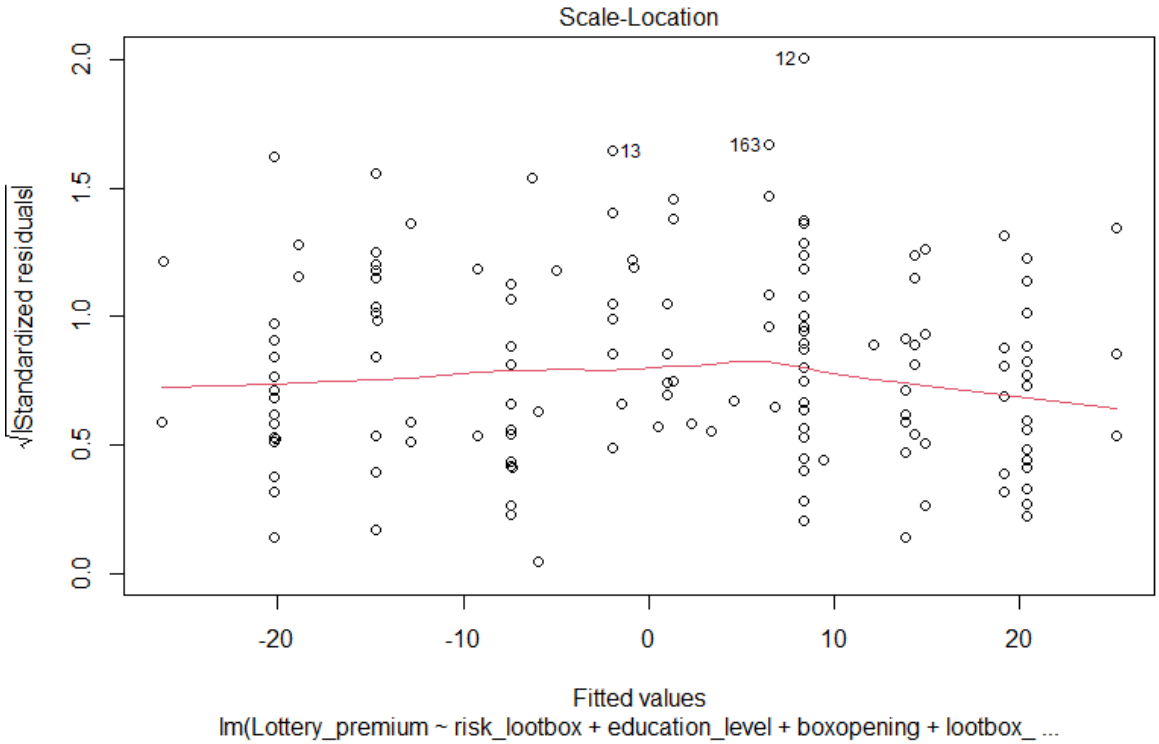

Scale-Location
√|Standardized residuals|
Fitted values
lm(amount_spent ~ age + casino_games + odds_search + Friends_open + box_ite ...
155
49
142
0.0
0.5
1.0
1.5
0
5
10
15
20
25
30

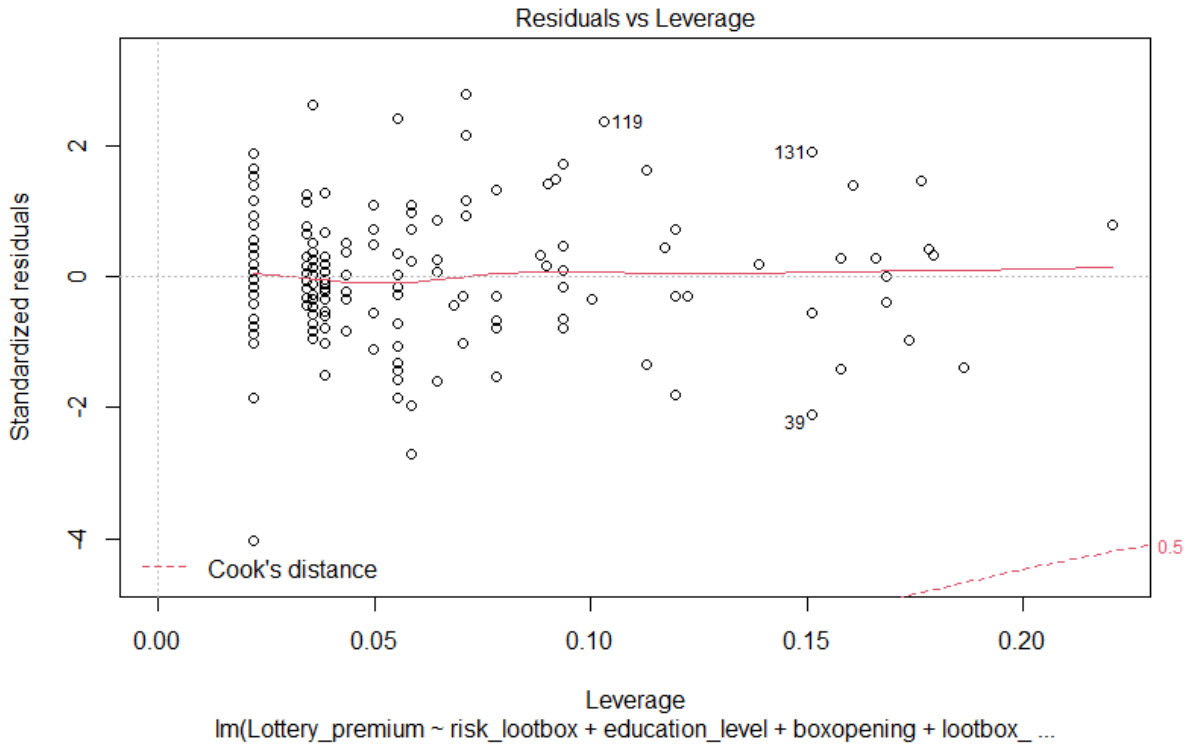

Residuals vs Leverage
Standardized residuals
Leverage
lm(amount_spent ~ age + casino_games + odds_search + Friends_open + box_ite ...
Cook's distance
145
159
142
-2
-1
0
1
2
0.00
0.05
0.10
0.15

Plots generated by R from linear regression for lottery premium






Scale-Location
√|Standardized residuals|
Fitted values
lm(Lottery_premium ~ risk_lootbox + education_level + boxopening + lootbox_ ...
12
13
163


Residuals vs Leverage
Standardized residuals
Leverage
lm(Lottery_premium ~ risk_lootbox + education_level + boxopening + lootbox_ ...
Cook's distance
0.5
119
131
39

Appendix B: Additional Regressions and tests

Table B.1: OLS regression for amount spent including all variables related to hypotheses

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	-3.176	1.963	-1.618	0.108
Age	0.198	0.075	2.646	0.009 **
casino_gamesYes	2.728	0.730	3.735	0.000 ***
odds_searchYes	-2.222	0.627	-3.545	0.001 ***
Friends_open	0.419	0.166	2.524	0.013 *
box_item_value_root	0.634	0.085	7.502	4.47e-12 ***
disposable_inc	0.005	0.001	5.402	2.42e-07 ***
Lottery_premium	0.186	0.022	8.297	4.72e-14 ***
hours_played	0.048	0.036	1.326	0.187
account_soldYes	1.084	2.027	0.535	0.594
box_openingYes	-0.303	0.751	-0.404	0.687

Residual standard error: 3.759 on 156 degrees of freedom

Multiple R-squared: 0.7594, Adjusted R-squared: 0.7455

F-statistic: 54.71 on 9 and 156 DF, p-value: < 2.2e-16

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

B-P test for amount_spent regression

studentized Breusch-Pagan test

data: model1

BP = 10.298, df = 7, p-value = 0.1723

B-P test for lottery_premium regression

studentized Breusch-Pagan test

data: model2

BP = 1.1854, df = 9, p-value = 0.9989

VIF test for amount_spent regression

Variable	VIF
age	1.64
casino_games	1.47
odds_search	1.02
Friends_open	1.12
box_item_value_root	1.13
disposable_inc	1.66
Lottery_premium	1.41

VIF test for lottery_premium regression

Variable	GVIF	Df	GVIF^(1/(2*Df))
Risk_lootbox	1.52	3	1,07
education_level	2,79	4	1,14
boxopening	1,63	1	1,27
Lootbox_pur_freq_month	1,40	1	1,19