VSB-TECHNICAL UNIVERSITY OF OSTRAVA

FACULTY OF ECONOMICS

Doctoral Dissertation Supplement

DISPOSITION EFFECT IN THE CONTEXT OF EXPERIMENTAL TRADING ON HIGHLY LIQUID FINANCIAL MARKETS

Field of study: Finance

Hana Dvořáčková

Ostrava, August, 2020

VSB-TECHNICAL UNIVERSITY OF OSTRAVA

FACULTY OF ECONOMICS

Hana Dvořáčková

Doctoral Dissertation Supplement

DISPOSITION EFFECT IN THE CONTEXT OF EXPERIMENTAL TRADING ON HIGHLY LIQUID FINANCIAL MARKETS

Study Program: Finance

Ostrava, August, 2020

Author of doctoral dissertation: Hana Dvořáčková

Department of Finance, EkF, VSB-TU Ostrava

Supervisor of doctoral dissertation: prof. Ing. Tomáš Tichý, Ph.D.

Department of Finance, EkF, VSB-TU Ostrava

The defense of the doctoral dissertation will be held before a committee for the defense of doctoral dissertation in the field of study finance at Faculty of Economics, VSB-TU Ostrava, Sokolská třída. 33, room E 308

The doctoral dissertation can be consulted at the Department of doctoral studies, Havlíčkovo nábřeží 38a, Room E 410

Contents

1 Introduction

The disposition effect has been described in the stock-investing context as a behavioral tendency of investors to hold on losing stocks for too long and sell winning stocks too early. This is based on Kahneman et al. (1979), the prospect theory to investment. According to the authors, "*a person who has not made peace with his losses tends to accept risks and gamble, which would not be otherwise acceptable.*" Later, Shefrin et al. (1985) published the first paper naming this behavioral effect as the disposition effect. According to Odean (1998), the most obvious explanations, i.e. explanations based on informed trading, re-balancing, or transaction costs, fail to capture essential features of the data. For the purpose of the said paper, records of 10 000 accounts at a large discount brokerage house were analysed to prove the investors' tendency to hold losing investments for too long and sell winning investments too early.

Several studies focused on the trader's behavior and the disposition effect were published over time. For instance, Barberis et al. (2009) investigated whether the prospect theory preferences can predict a disposition effect, with Kaustia (2013) including a chapter focused on the disposition effect in Behavioral Finance: Investors, Corporations, and Markets. These authors also found strong evidence for the disposition effect and explained it in terms of investor characteristics. Chen et al. (2007) studied the investment decision making in an emerging market. According to them, Chinese investors, among others, tend to sell stocks that have appreciated in price, but not those that have depreciated in price, which is consistent with the disposition effect of acknowledging gains but not losses. Results of Marciukaityte et al. (2012) are in accordance with the statement that the disposition effect increases the supply of winning stocks and depresses their prices. Contrary, Locke et al. (2005) found no evidence of any contemporaneous measurable costs associated with the disposition effect. The open questions are: How do act students, who are trading under the same conditions as real traders (trading platform, transaction costs, real-time movements of financial markets), but do not bear the risk of real losses or gains? Is the disposition effect such a natural behavioral bias that will be present also in this kind of experimental data? Answers (not solely) to these questions were answered in the thesis.

2 Goal and Structure of the Thesis

The disposition effect is a crucial phenomenon, which might lead even to global financial market disorders. The thesis aims to examine whether the disposition effect can be identified in selected experimental trading data and if so, what are the affecting factors and what are the differences within various groups of traders.

The first, theoretical part of the thesis focuses on the theory related to the topic. Specifically, Chapter 2 focuses on general background of decision making, financial markets, and the OANDA trading platform. Chapter 3 contains the methodology necessary to fulfill the aim, as for example, relevant statistical methods, GUHA method, or logistic regression analysis. Moreover, the *TimeLtoP* ratio used for assessing the potential disposition effect is explained in this chapter.

The second part of the thesis focuses on the data description and results and their interpretation. Besides other variables, the gender difference in approach to taking risk or the tendency to feel the addiction to trading was tested. The aim of the thesis was achieved by examining the essential hypotheses and hypotheses generated by the GUHA method. The use of experimental data makes this thesis significantly different from works of previous research, which, to author's best knowledge, were based on data collected from real investors [1] for which specific characteristics and standardized behavioral biases, such as the self-selection bias (Adams et al. (2012)) can be assumed, or based on the experiment where traders were "trading" with randomly modelled stocks (Weber et al. (1998)).

Research based on experimental data has its disadvantages, as well as advantages. The distinct disadvantage is that students were not trading with real money, therefore, they might have acted differently than they would do in reality. On the other hand, this disadvantage can easily become an advantage, as students were not affected by their wealth or attitude to handling money. If someone trades one hundred thousand USD, his decisions are definitely affected by the share of this amount on his total wealth. The advantages and disadvantages are discussed in depth in Subchapter 3.5 of the thesis, together with the problematics of motivation to achieve the best result possible.

Regarding hypotheses stated by the author, first of all, it was examined if the disposition effect

[1] For example, Shefrin et al. (1985), Odean (1998), Lehenkari (2012), Barberis et al. (2009) Cheng et al. (2013), Eom (2018), Muhl et al. (2018), Sarmiento et al. (2019).

is present at all in the data set; hence the length of loss trades is significantly longer than length of profitable trades. Moreover, the effect of gender, size of the trade, and culture was tested.

Consequently, the GUHA method, which is suitable for automatic generation of hypotheses within large data sets[2], was applied on two versions of the data set, containing ten explanatory variables and one dependent variable - length of trades. Firstly, the original data set containing over ten thousand observations was used to explore all relevant and interesting hypotheses of trades' assignment to a given length interval. Those with the highest *Lift* were chosen and statistically tested. Secondly, the data was aggregated on the level of particular student to avoid inappropriate multiplication of some variables (mostly answers from the questionnaire questions) and again the GUHA was applied.

The last part of the empirical chapter was focused on the logistic regression analysis. This method was used to asses the impact of several variables on the dependent variable called "*DE*", which equals to 1 if the *TimeLtoP* is higher than 1.1, otherwise the *DE* is equal to 0. The variables, whose influence was tested are, for example, gender, risk attitude, trades' size, or tendency to addiction.

[2] See Hájek; Havel, et al. (1966), Hájek; Holeňa, et al. (2010) Hájek (2003), Pudlák et al. (1979), Turunen (2009).

3 Content of the Thesis

4 Methodology

Within this section two main groups of methodology are described. Firstly the methods used for testing of hypotheses (general statistical methods), generating new hypotheses (GUHA), and modelling (logistic regression. The second group is represented by methodology of data collection.

4.1 Mathematical Methods

The thesis is based on three mathematical methods. Firstly the general statistical methods for hypotheses testing, particularly the F-test and t-test were used. The author will not focused on the detailed description of this methodology, as it is generally well known.

The next, less known method is GUHA, which is the acronym for the General Unary Hypotheses Automaton. This method was firstly described in Hájek; Havel, et al. (1966) and it is one of the oldest data mining methods used for generating of systematical hypotheses supported by the given empirical data. According to Turunen (2009), the GUHA method is pursuant to the well-defined first order monadic logic, which contains generalized quantifiers of finite models. A GUHA procedure generates statements on association between complex Boolean attributes, which are built from the predicates corresponding to the columns of the data matrix described hereinafter.

This method is essentially suitable for an exploratory analysis (having no specific hypothesis to be tested, only a general problem) of large data, typically used, for instance, in pharmaceutical and medical research, the author is not aware of previous application of this method in the field of behavioral finance or financial markets. The data form a rectangle matrix of zeros and ones, rows corresponding to objects belonging to the sample and columns corresponding to the researched variable (attributes). Typically the matrix has hundreds to thousands of rows and dozens of columns. As stated in Hájek; Holeňa, et al. (2010), let's assume to $P_1, \dots, P_n$ to be names of the attributes. For each $P_i, \neg P_i$ is its negation. An elementary conjunction of length $k (1 \leq k \leq n)$ is a conjunction of k literals in which each predicate occurs at least once, for example $\neg P_3;\ P_1 \wedge \neg P_3 \wedge P_7$. Similarly to this, an elementary disjunction is defined as, for example, $P_1 \vee \neg P_3 \vee P_7$. An object satisfies an elementary conjunction in case of satisfaction of all its members and it satisfies an elementary disjunction if it satisfies at least one of its members.

Let $0 \leq P \leq A$; formula $A \Rightarrow S_p$, where A is an Antecedent, i.e. an elementary conjunction and S is a Succedent, i.e. an elementary disjunction, is true in given data set if at least 100p percent of objects satisfying A satisfies also S, i.e. $\frac{a}{r} \geq p$, where r is the number of objects satisfying A and a is the number of objects satisfying A together with S. The Antecedent A is *t-good* if at least t objects satisfy it.

The software used for data mining using GUHA method is called the *LISp-Miner* and it has been developed in 1966 at the Faculty of Informatics and Statistics of University of Economics in Prague. The software provides users with six GUHA procedures. GUHA procedure is a computer program which based on imputed data provides a simple definition of relevant, thus potentially interesting patterns. A GUHA procedure generates particular relevant patterns and tests whether they are true within analyzed data. Afterwards the output contains all prime patterns which are true in analyzed data and do not immediately follow from the other, more simple output pattern.

Pudlák et al. (1979) focused on the problematic of mechanised hypotheses formation, they proposed: Let M stand for a data matrix (two-valued); call an elementary disjunction positive if it contains no negative literals, such a disjunction is true in M if it is satisfied by all objects. They state

Theorem 4.1

(A) The set of all pairs (M,k) such that M is a data matrix and there is a positive elementary disjunction of length $\leq k$ true in M is NP-complete [3].

(B) The set of all pairs (M,k) such that M is a data matrix and there is an arbitrary elementary disjunction of length $\leq k$ in M is deterministically decidable in the time 2^{log^2}.

Another hypotheses, focused on the form $\varphi_i \sim \psi_i$, where $\sim$ represents a given generalized quantifier, are particular formal sentences and it is possible to combine them using connectives of the classical logic, giving a new sentence called a sentence in normal form. This sentence is a tautology of the quantifier $\sim$ if it is true in each data matrix (again two-valued). It is an implicational tautology

[3]NP is an acronym for non-deterministic polynomial time, which is a complexity class used to classify decision problems. A set X of formulas is NP-Complete if the set is in NP and each Y in NP is reducible to X in polynomial time, see Hájek (2003), or Goldreich (2010).

in case of being true in each data for any implicational quantifier $\sim$. Furthermore, for associational tautology and $\mathcal{K}$-tautology for any other class of quantifiers. As an example see subsequent formula:

$$(\varphi \sim_{\alpha}^{\text{Fish}} \psi) \rightarrow (\psi \sim_{\alpha}^{\text{Fish}} \varphi), \tag{4.1}$$

where $\sim_{\alpha}^{\text{Fish}}$ is the Fisher quantifier with significance α, which is a tautology of this quantifier. According to this formula the Fisher quantifier is symmetric. The formula:

$$[(\varphi \,\&\, p) \implies \xi] \rightarrow [\varphi \implies (\xi \vee \neg p)] \tag{4.2}$$

represents an implicational tautology. Tautologies can be used to deduce new hypotheses from hypotheses once found.

Theorem 4.2

(A) The set of all formulas in normal form that are implicational tautologies is co-NP-complete [4].

(B) Analogously for associational tautologies.

All GUHA procedures provided by the *LISp-Miner* work with data matrices, as an example is presented Figure 4.1, a data matrix $\mathcal{M}$. Each row of the matrix corresponds to one observed object and each column corresponds to one attribute. Each attribute has a finite number of *categories*, i.e. possible values. Due to this fact, these attributes are called *categorial attributes*. Attributes A_1 and A_{100} are examples of categorial attributes.

Individual patterns carry categorial attributes, Boolean attributes, or both. A Boolean attribute is created using logical connectives from the basic Boolean attributes. According to Hájek; Holeňa, et al. (2010), the basic Boolean attribute is the expression $A(\alpha)$, where $\alpha \subset \{a_1, \dots, a_k\}$ and $\{a_1, \dots, a_k\}$ is the set of all categories of the attribute A. A basic Boolean attribute $A(\alpha)$ is true in the row o. The given α is the coefficient of the literal $A(\alpha)$. Expressions $A_1(1, 2)$ and $\neg A_{100}(6)$ are examples of literals, as shows Figure 4.1.

[4] Co-NP is a complexity class, a set X of formulas is in co-NP if and only if its complement is in NP. A set X of formulas is co-NP-complete if it is in co-NP and if each Y in co-NP is polynomial-time many-one reducible to it. Each co-NP-complete problem is a part of an NP-complete problem (see Hájek (2003), or Goldreich (2010)). Equality of classes NP and co-NP is one of unsolved problems in computer science.

Figure 4.1: Data Matrix $\mathcal{M}$

object i.e. row of $\mathcal{M}$	columns of $\mathcal{M}$ i.e. attributes				examples of literals		cards of categories of attribute A_1		
	A_1	A_2	$\dots$	A_{100}	$A_1(1,2)$	$\neg A_{100}(6)$	$A_1[1]$	$A_1[2]$	$A_1[3]$
o_1	1	7	$\dots$	4	T	T	1	0	0
o_2	3	4	$\dots$	6	F	F	0	0	1
o_3	2	9	$\dots$	9	T	T	0	1	0
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
o_n	3	1	$\dots$	6	F	F	0	0	1

Hájek; Holeňa, et al. (2010)

There are three types of contingency tables to verify all patterns, all of them presents Figure 4.2:

- $4ft - table$, $4ft(\varphi,\ \psi,\ \mathcal{M})$ of the Boolean attributes φ and ψ in $\mathcal{M}$;
- $KL - table$, $KL(R,\ C,\ \mathcal{M})$ of the categorial attributes R and C in $\mathcal{M}$;
- $CF - table$, $CF(R,\ \mathcal{M})$ of the categorial attribute R in $\mathcal{M}$.

Figure 4.2: Matrices $4ft(\varphi,\psi,\mathcal{M}), KL(R,C,\mathcal{M})$, and $CF(R,\mathcal{M})$

$\mathcal{M}$	ψ	$\neg\psi$
φ	a	b
$\neg\varphi$	c	d

$4ft(\varphi,\psi,\mathcal{M})$

$\mathcal{M}$	c_1	$\dots$	c_L
r_1	$n_{1,1}$	$\dots$	$n_{1,L}$
$\vdots$	$\vdots$	$\ddots$	$\vdots$
r_K	$n_{K,1}$	$\dots$	$n_{K,L}$

$KL(R,C,\mathcal{M})$

$\mathcal{M}$	r_1	$\dots$	r_K
	n_1	$\dots$	n_K

$CF(R,\mathcal{M})$

Hájek; Holeňa, et al. (2010)

The 4ft table is a quadruple $\langle a,b,c,d\rangle$ of natural numbers, whereas a represents number of rows of $\mathcal{M}$ satisfying φ and also ψ, b represents number of rows satisfying φ but not ψ, c represents number of rows satisfying ψ but not φ, and c represents number of rows not satisfying φ nor ψ. Furthermore, it is supposed that the attribute R has the categories $r_1,\dots,r_K$ and the attribute c has the categories $c_1,\dots,c_L$. The expression $n_{k,l}$ in the KL-table shows the number of rows of the data matrix $\mathcal{M}$ for which the value of attribute R is equal to r_k and the value of attribute C is equal to c_l. Finally the expression n_k in the CF-table $CF(R,\mathcal{M}) = \langle n_1,\dots,n_K\rangle$ shows the number of rows in data matrix $\mathcal{M}$ for which the value of attribute R is r_k.

All contingency tables are computed based on representation of analysed data by so-called cards of categories. The card of categories is a string of bits. For instance, the attribute A_1 of the data matrix $\mathcal{M}$ presented in Figure 4.1 has three categories $\{1, 2, 3\}$, which means that the attribute A_1 is represented by cards $A_1[1], A_1[2], A_1[3]$ of categories 1, 2, and 3 respectively. The card $A_1[1]$ of category 1 is a string of bits. Every row of given matrix $\mathcal{M}$ corresponds to one bit of the card $A_1[1]$. There is "1" in the bit corresponding to the row o_i if and only if o_i row is of value 1. The same applies for the other categories and attributes, see Figure 4.1.

Above described cards of categories are used to build corresponding cards of Boolean attributes and to compute particular frequencies from contingency tables. Card $C(\varphi)$ of Boolean attribute φ is a string of bits. Each row of the data matrix serve as one bit of $C(\varphi)$. There is "1" in the i^{th} bit only if φ is true in row o_i. Hence it is obvious that $C(\varphi \wedge \psi) = C(\varphi) \dot{\wedge} C(\psi)$, $C(\varphi \vee \psi) = C(\varphi) \dot{\vee} C(\psi)$, and $C(\neg\varphi) = \dot{\neg} C(\varphi)$. Here $C(\varphi) \dot{\vee} C(\psi)$ represents a bit-wise conjunction of bit strings $C(\varphi)$ and $C(\psi)$, analogously for $\dot{\vee}$ and $\dot{\neg}$. Furthermore it applies that $C(A_1(1, 2)) = A_1[1] \dot{\vee} A_1[2]$ for basic Boolean attribute $A1(1, 2)$, etc.

As already mentioned, currently there are six GUHA procedures available in the *LISp-Miner*: 4ft-Miner, KL-Miner, CF-Miner, SD4-Miner, SDKL-Miner, and SDCF-Miner.

The **4ft-Miner** looks for association rules $\varphi \approx \psi/\gamma$ which means that the Boolean attributes φ and ψ are associated in the way given by the symbol $\approx$, called *4ft-quantifier*. The 4ft-quantifier represents a special case of the generalized quantifier, its semantic is given by a function associating with each 4ft-table the truth value 0 or 1. The association rule $\varphi \approx \psi/\gamma$ is true in the matrix $\mathcal{M}$ if the relevant condition is satisfied in the 4ft table $4ft(\varphi, \psi, \mathcal{M})$. The quantifier $\implies_{p,t}$. Further example of the 4ft-quantifier is the 4ft-quantifier $\sim^{+}_{p,t}$ of above average dependence, defined for $0 < p$ and $t > 0$ by the condition $\frac{a}{a+b} \geq (1 + p)\frac{a+c}{a+b+c+d} \wedge a \geq t$. Accordingly, the relative frequency of the objects satisfying ψ within objects satisfying φ is at least 100p percent higher than the relative frequency of the objects satisfying ψ within the whole matrix $\mathcal{M}$ and that there are at least t objects satisfying both φ and ψ. Totally there are 17 types of the 4ft-quantifiers implemented in this procedure.

The **KL-Miner** looks for KL-patterns $R \sim C/\gamma$, which means that frequencies of categories of attribute R satisfy the condition given by $\sim$ when another condition given by the Boolean attribute γ is satisfied. In this case, the symbol $\sim$ is called *KL-quantifier*. A KL-quantifier corresponds to a condition concerning the KL-table $KL(R, C, \mathcal{M}')$ of the categorial attributes R and C in the matrix $\mathcal{M}'$. The KL-pattern is true within the matrix $\mathcal{M}$ if the condition represented by $\sim$ is satisfied within the KL-table $KL(R, C, \mathcal{M}/\gamma)$ of the categorial attributes R and C in the matrix $\mathcal{M}/\gamma$.

The **CF-Miner** looks for CF-patterns in the form of $\sim R/\gamma$. The pattern means that frequencies of categories of attribute R satisfy the condition given by $\sim$, the *CF-quantifier*, whereas other condition given by the Boolean attribute γ is met. The CF-quantifier corresponds to a condition regarding the CF-table $CF(R, \mathcal{M}')$ of the categorial attribute R in the matrix $\mathcal{M}'$. The CF-pattern is true in case of satisfaction of the condition given by $\sim$ in the CF-table $CF(R, \mathcal{M}/\gamma)$ of the categorial attribute R in the data matrix $\mathcal{M}/\gamma$.

The **SD4ft-Miner** looks for SD4ft-patterns of the form $\alpha \bowtie \beta : \varphi \approx^{SD} \psi/\gamma$, whereas α, β, γ, φ, and ψ represent Boolean attributes derived from the columns of analyzed data matrix $\mathcal{M}$. The attributes α and β define two subsets of rows, the attribute γ defines a condition. Finally, the attributes φ and ψ are *antecedent* and *succedent* of a given association rule $\varphi \approx \psi$. The SD4ft pattern means the the subtest given by the Boolean attributes α and β differs as for validity of association rule $\varphi \approx \psi$ when the condition given by γ is met. The symbol $\approx^{SD}$ is called *SD4ft-quantifier* and represents a measure of difference. The SD4ft-quantifier corresponds to a condition regarding two 4ft-tables $\langle a, b, c, d\rangle$ and $\langle a', b', c', d'\rangle$. The SD4ft-pattern $\alpha \bowtie \beta : \varphi \approx^{SD} \psi/\gamma$ is true in the data matrix $\mathcal{M}$ if the condition corresponding to $\approx^{SD}$ is met for the 4ft-tables $4ft(\varphi, \psi, \mathcal{M}/(\alpha \wedge \gamma))$ and $4ft(\varphi, \psi, \mathcal{M}/(\beta \wedge \gamma))$.

The procedures **SDKL-Miner** and **SDCF-Miner** are derived from the procedures KL-Miner and CF-Miner in the similar way as the procedure SD4ft-Miner is derived from the 4ft-Miner procedure.

By GUHA all hypotheses interesting from the point of view of a given general problem are systematically created on the base of particular data, while the output is a list of all prime patterns. According to Hájek; Havel, et al. (1966), the main principle is that all interesting hypotheses contains a dilemma as "all" means most possible whereas "only interesting" means "not too many".

However, the GUHA provides polyfactorial hypotheses, therefore not only hypotheses relating one variable with another but the method is expressing variables among single variables, pairs, triples etc. Resulting hypotheses are created but not statistically tested, hence it is necessary to make further statistical tests.

When using this method it is necessary to understand terms *Support*, *Confidence*, *Lift*, and *Count*, also used in Chapter 5 of the thesis. *Support* represents the percentage of rows from the overall data set satisfying the left side of hypotheses generated by GUHA, whereas *Count* represents this information in absolute value. Variable *Confidence* shows the percentage of rows to which the right side of hypotheses is applicable. Variable *Lift* defines the increase of *Confidence* of data filtered based on the GUHA results comparing to the overall data set.

The last method used, as well as described in the Thesis is the logistic regression, which is according to Hosmer; Lemeshow (2004) a significant part of the data analysis, testing the relationship between the dependent variable and at least one independent variable. In general, the logistic regression method has two main goals:

1. to find independent variables contributing to the explanation of dependent variables;
2. to estimate a logistic regression model, which will be used for group classification.

Logistic Regression Model Estimation

An odds ration compares the probability of a phenomenon occurring and not occurring, represented by formula:

$$\frac{\pi}{1-\pi}, \tag{4.3}$$

where π is probability of phenomenon occurring. By logarithmization of previous formula a value on the interval $(-\infty, \infty)$ is obtained. After that the logistic model corresponds to the equation:

$$\pi(x) = \frac{e^{\beta_0+\beta_1 x_1+\cdots+\beta_p x_x}}{1+e^{\beta_0+\beta_1 x_1+\cdots+\beta_p x_p}}. \tag{4.4}$$

By the logit transformation with the domain $(-\infty, \infty)$, $\pi(x)$ may be transformed to logit as follows:

$$g(x) = \ln\left[\frac{\pi(x)}{1-\pi(x)}\right] = \beta_0 + \beta_1 x_1 + \cdots + \beta_p x_p. \tag{4.5}$$

Maximum Likelihood Estimation

Unlike the linear regression using the least squares method, the logistic regression coefficients are usually estimated by using the maximum likelihood estimation, which is based on a dichotomous dependent variable. If Y is coded as 0 or 1, there is $P(Y = 1/x)$ for probability that Y is equal to 1 with given x. Similarly $P(Y = 0/x)$ refers to the probability that Y is equal to 0 with given x. The contribution to the likelihood function is $1 - \pi(x_i)$, where the quantity $\pi(x_i)$ denotes the value of the $\pi(x)$ calculated at x_i. In the Hosmer; Lemeshow (2004) is presented following formula:

$$\pi(x_i)^{y_i} \left[1 - \pi(x_i)\right]^{1-y_i} . \tag{4.6}$$

Assuming independence of observations, the likelihood is a function of the response variable in case of binominal distribution stated as follows:

$$l(\beta) = \prod_{i=1}^{n} \pi(x_i)_i^y \left[1 - \pi(x_i)\right]^{1-y} . \tag{4.7}$$

By logarithmization of the previous formula according to Hosmer; Lemeshow (2004), the subsequent formula is obtained:

$$L(\beta) = \ln\left[l(\beta)\right] = \sum_{i=1}^{n} \{y_i \ln\left[\pi(x_i)\right] + (1 - y_i) \ln\left[1 - \pi(x_i)\right]\}. \tag{4.8}$$

To find value β while maximizing $L(\beta)$, $L(\beta)$ is differentiated with respect to β_0 and β_1, final likelihood equations are laid to zero:

$$\sum \left[y_i - \pi(x_i)\right] = 0, \tag{4.9}$$

$$\sum x_i \left[y_i - \pi(x_i)\right] = 0. \tag{4.10}$$

The final general logistic model is of subsequent form:

$$\ln\left(\frac{L_{(1)}}{L_{(0)}}\right) = b_0 + b_1 x_1 + b_2 x_2 + \cdots + b_p x_p, \tag{4.11}$$

where $\ln\left(\frac{L_{(1)}}{L(0)}\right)$ is logarithm of the probability ratio (logit) and $b_0, \ldots, b_p$ are estimates of regression coefficients.

To estimate coefficients the IBM/SPSS software will be used. There are two methods available - ENTER and STEPWISE, the main difference is in the method of predictors input. Using the ENTER method, all predictors are imputed together, followed by calculation of parameters' approximation. Contrary using the STEPWISE method, according to Field (2013), predictors are added gradually.

The logistic regression model needs to be statistically verified to evaluate the predictive ability of estimated model. Following methods were used in the thesis:

- **-2 Log likelihood** (-2LL) ratio test evaluates the difference in the natural logarithm of the partial likelihood functions, corresponding to the alternative and the null hypothesis;
- **Cox & Snell $\mathbf{R}^2$** expresses the predictive ability of a model, specifically how many variations of dependent variable can be explained;
- **Wald chi-square test**, which is based on the maximum likelihood approximation of the first component from the estimated coefficient;
- **Hosmer–Lemeshow test**, a goodness-of-fit test;
- **Receiving Operating Characteristic Curve** also called the ROC curve, is a graphical, analytical tool, depicting the diagnostic ability of a binary classifier system;
- **Gini Coefficient**, also used for test of the goodness-of-fit of a model.

4.2 Data Collection Methodology

The data set used in the thesis has been collected by Jochec and Dvořáčková (Dvorackova et al. (2018)) from 2009 to 2015 during lectures in several countries (for instance, the USA, New Zealand, Kazakhstan, etc.), moreover in 2016 at the Faculty of Economics of the VSB-TUO. In total around 300 students were trading on the OANDA FX Trade Practice platform, described in the previous Chapter 6.1, with currency pairs and CFDs.

Initially students were given \$100 000, meanwhile the trading period was standardized and took three months. As those students did not trade with real money, they were motivated to achieve as good result as possible by a financial reward together with some extra points for the exam on account of the winner (student with the highest account balance at the end of the trading period). One of the learning objective was to experience the first-hand trading and use explained technical analysis techniques in practise. At the end of the given trading period students had to submit in detail recorded trading history together with their answers to a short questionnaire and demographic information. There were collected also students' login information and passwords of their official game account therefore it was not possible to change the account later, reset losses, use more accounts, etc.

The data generation and collection process proceeded as follows: The course was started with a series of lectures and assignments designed to explain the currency trading basics and the use of the trading platform. The game was launched sometime between the second and fourth week of semester, and was running for the rest of the semester (60 to 90 days). Soon the focus shifted on another topics and the game continued in the background. The rules and the interference with students were minimal, students were not asked for any specific strategy, neither encouraged nor discouraged to the use of fundamental or technical analysis; there was no "preferable" amount, frequency, size, or currency to be traded. The winner was the student with the highest trading account balance at the end of the given time period. The ending profit or loss did not affect the course grade except that the winner (and only the winner) earned few extra points towards the final course grade, and in some cases a voucher to a bookshop.

The setting of the game in the currency market is convenient, currency market is liquid, see Mancini et al. (2013), and close to efficient, which was already discussed in Subchapter ě.č of the thesis. It is difficult to make meaningful price predictions and trading is more a matter of luck then skill. Accordingly, the skill component does not distort the picture and the trading patterns and strategies tend to be more behavioral in nature. The experimental setting has obvious disadvantage, the money is not real, thus the joy of winning (pain of losing) is moderated. This should be slightly counteracted by awarding the winner to support students' motivation to achieve as good result as possible. The counterargument might be that the "winner takes it all" reward scheme is problematic, there is no incentive for scoring second (third etc.); similarly, scoring low does not bear any penalty. This and the fact that it is hard to predict currency rates even for professional traders means that students were in effect encouraged taking higher-than-normal risk and engaged in "all or nothing" gamble. It was not possible to perfectly rule those problems out, however, there is no indication of more frequent occurrence of large bets on the last few days of the game, which would point out a tendency towards a pure gambling. It can be assumed that students derived some benefits also from simply doing well, even if not the best. This could result from the long-term continuance of the experiment and the psychological benefit (cost) of favourable (unfavourable) comparison with the peers. It is possible to see a parable with gaming. People around the world are playing several online games which are for free and provide no reward for playing, just comparison with other gamers within a ranking. According to Stanley (2012) the social media gaming has around 800 million monthly users around the world, thus 800 million of people are motivated to play a game every month to keep a peer pride among people who even do not know. They are even willing to pay for some extra features, to achieve better position, see Gainsbury et al. (2017).

This conjecture is supported by student comments and informal feedback. The assumption is that in spite of the singular incentive, the students chose trading strategies without trying to "game the system" or engaging in an ultimate all-in gamble. The long-term continuance of the research can b taken as another disadvantage. Students were trading on real trading platform, thus the real global economic situation was affecting their decision making. On the other hand, a potential fear from some bad development of the economical situation could be moderated by the fact, that they did not trade with real money, thus it is supposed that their decisions were highly natural. Furthermore, economical cycles are, and always be present, thus it should not be totally removed.

One can identify various advantages of the experimental setting. Firstly, participants do not self-select into roles, thus the sample is less biased: let's take, for example, the effect of gender on trading decisions. To compare behavior of actual female and male traders is problematic because female traders might have some male characteristics, which made them to select the trading profession and helped them to succeed in it (self-selection and survivor-ship bias). As proved in Adams et al. (2012), females who are working in top management positions tend to overtake male behavior characteristics and, besides others, show different approach to the risk than general females. This might be related also to the highly demanding work in the trading industry. The sample of students/traders does not represent the population ideally because all of them decided to study business, but not all of them would like to do the trading career.

Second advantage of the experimental setting is the homogeneity of objectives. In real situation, someone might set up a currency position as a hedge for some other asset. For example, if somebody's savings are predominantly in EUR but expected expenses in USD, that person might want to open a short position in EUR/USD in order to hedge the EUR exposure of the savings account. The loss on currency position offsets the gain on the savings account (and vice versa), consequently decreasing the volatility of wealth. In case that the real traders' behavior is analysed, their goals are not known so the question is, are they speculating or hedging? Different objectives would lead to different trading strategies. Moreover, in the real life, trader's wealth is given by the sum of different assets, hence a loss in OANDA may be compensated by gain of another asset and a high net asset value person may trade differently from a low net asset value person.

In spite of not being offered monetary compensation, the data shows that students took the trading game seriously. It is supposed that it was partly caused by the fact that students found the game interesting, as shown by the questionnaire result, 75% of students answered that they traded because it was a course requirement and also interesting for them, 5% traded because it was purely interesting, and 20% of students traded only because it was a course requirement. According to another question, 44% of students actually had at least some feelings of addiction during the game.

Students were strongly self-motivated partly because they played the game while being introduced to the world of international finance as the course progressed (thus seeing its relevance and connection with the real world). The other part of their motivation could have been the peer pride. Students were often discussing their trading success and failures or boastfully showing others their impressive results on their OANDA – enabled smart phones. The fact that they competed with and bench-marked their results against their classmates and friends over an extended period possibly made the game more interesting (compared to some short laboratory experiment with strangers).

For simplification of calculations a special variable named $TimeLtoP$ was specified. It is calculated as shown below in formula (4.12):

$$TimeLtoP = \frac{\text{average holding period of unprofitable trades}}{\text{average holding period of profitable trades}} \tag{4.12}$$

The resulting ration provides us with information about potential presence of the disposition effect. In case of the result equal to 1 there is no disposition effect, the higher $TimeLtoP$, the higher probability of the disposition effect presence. Additionally, the ration less than 1 shows the reverse disposition effect. For example, $TimeLtoP$ equal to 2 can be interpreted as an average holding period of unprofitable trades two times longer than average holding period of profitable trades.

Odean (1998), Weber et al. (1998), or Sarmiento et al. (2019) used the frequency approach to test the disposition effect, however, in the thesis the duration approach is preferred. The duration approach is advantageous in high frequency trading environment, herein represented by FX market, as traders are likely to hold their positions for a short term and make more trades, see Eom (2018).

5 Results and Conclusion of the Thesis

The results of the thesis are divided into three groups. Firstly, the initial hypotheses presented by the author are presented, followed by the hypotheses generated by the GUHA Method. Afterwars, the logistic regression results are described. The very last subchapter is the Conclusion.

5.1 Essential Hypotheses Stated by the Author

Within Subchapter 4.1 of the thesis four hypotheses stated by the author were tested. All of them were proceeded under the same methodology, i.e. application of the F-test to asses the (in)equality of variances was followed by the relevant two-tailed t-test to asses difference in length of trades means. Afterwards, both particular groups were tested in the similar way, using the one-tailed t-test to see if the mean length of loss trades is higher than mean length of profitable trades. In case of insignificant results, also the two-tailed t-test was performed. The t-tests were performed on the hypotheses from several points of view.

The first hypothesis assumed the disposition effect in the data set, the F-test resulted in rejection of the null hypothesis, hence expecting the variances inequality. According to the one-tailed t-test for inequal variances, mean length of profitable trades is smaller than mean length of unprofitable trades, which supports the basic logic of the disposition effect. Herewith it is possible to presume the disposition effect presence in this particular data set. Afterwards also the *TimeLtoP* of profitable and loss trades was tested in the same way, whereas the mean *TimeLtoP* of profitable was confirmed to be smaller than that of loss trades. Based on this result it is possible to assume that even though students were trading on a demo account thus avoided the risk of loosing own funds, they succumbed to the same bias as traders using real money.

Secondly the influence of gender on the holding period was examined. Based on the F-test and the t-test results it is possible to assume statistically significant influence of gender on the length of trades. Additionally, the effect of profit and loss was examined. According to the one-tailed tests results, loss trades were held significantly longer than profitable trades made by females as well as by males. Another testing was performed to asses the influence of gender on the length of profitable and loss trades separately. Also in this case the difference was confirmed to be statistically significant.

The third hypothesis was built up on the belief that together with increasing trades' volume, thus potential loss, traders will face stronger disposition effect. According to given statistical tests it is possible to state that size of trade has an effect on the length of trade. In this case, the general hypothesis was also supported by tests focused on the impact of profit and loss with respect to the trades volume. The one-tailed t-test confirmed that small loss trades were held significantly longer than small profitable trades, but for large trades all statistical tests were found insignificant. Also in this case the difference for profitable and loss trades separately was examined. The difference between small and large trades from the point of view of profitable as well as loss trades was found statistically significant.

Final hypothesis assumes difference in holding periods of trades between students originating from New Zealand and from other countries, due to the general cultural differences. Statistical tests confirmed the difference statistically significant, students from New Zealand tend to keep trades significantly longer than other student, whereas their mean holding period was 49% higher. Moreover, one-tailed t-test confirmed that students from New Zealand held loss trades significantly longer than profitable ones. This was not possible to confirm for students from other countries. Final tests to asses solely profitable or loss trades also proved statistically significant difference between students from New Zealand and other countries.

5.2 Hypotheses generated by the GUHA Method

The GUHA method, generally used for the automated generation of hypotheses was used on two data sets to generate additional hypotheses. The explained variable was in all cases the *length of trade*, representing holding period of trades in days. Ten explanatory variables were chosen to generate hypotheses. First data set contained over ten thousand observations, whereas the explanatory variable (*length of trade*) was divided into three intervals and GUHA seems to be successful in the hypotheses generation in this case. From all generated hypotheses only those with the highest *Lift* measure were chosen. Based on the results, a loss making male, who did not feel addiction while trading, enjoys risk and has ever done adrenaline sports has more often shorter length of trade than others. The trader, who will probably have trades within the Interval A2 is defined by profit on the trade in the height within interval $\$\langle 50, 200\ 000\rangle$ and have never done any speculation. To the third Interval A3 will probably be included a trader who has never done any speculation, without any other definition. All mentioned hypotheses were found statistically significant.

The second data set was created by aggregation of the first one on the level of particular student, therefore it contained around three hundred observations, which is significantly lower than in previous case and the result was affected by this fact. Intervals were affected by extreme values, thus interval B1 and B3 did not contain enough data to achieve sufficient results. The only statistically significant hypotheses were generated for interval B2, whereas six of them had the same *Lift*. Based on obtained hypotheses, six combination of factors leading more often to the average length of trades within the Interval B2 were found:

1. a man, originating from New Zealand, who felt an addiction to trading and have ever made money by speculation;
2. a man, achieving average profit within the interval $\$\langle 50, 200\ 000\rangle$, who felt an addiction to trading and has ever done an adrenaline sport;
3. a man, originating from New Zealand, achieving average profit within the interval $\$\langle 50, 200\ 000\rangle$, who felt sometimes an addiction to trading and have ever made money by speculation;
4. a man, originating from New Zealand, with average volume of trades over median, who felt an addiction to trading and have ever made money by speculation;
5. a man, achieving average profit within the interval $\$\langle 50; 200\ 000\rangle$, with average volume of trades over median, who felt an addiction to trading and has ever done an adrenaline sport;
6. a man or a woman originating from New Zealand, profit maker at the end of trading period, who has never made money by speculation.

5.3 Logistic Regression Analysis

The last point of the empirical part was to explore how particular variables affect the disposition effect represented by the *TimeLtoP* ratio. Let's assume that students showing *TimeLtoP* below 1.1 has very small or no tendency to the disposition effect, all other students significantly succumbed to that behavioral bias. To asses the impact of particular explanatory variables a logistic regression model with the dependent variable *DE* (Disposition Effect) was created. Variable *DE* is equal to zero for traders showing *TimeLtoP* ≤ 1.1, otherwise is equal to one. The model is built up on 80% of the data set, thus 8 693 observations, and later tested on the rest of the data sample. To make the selection of observations, a sample of 2 010 pseudo-random numbers from the interval $\langle 1, 10\ 548\rangle$ was generated by the RAND function in the MS Excel. Observations were descendingly numbered and those involved in testing of the model were chosen according to the assigned number by the VLOOKUP function. The rest of observations were used for building up the model. The ratio of $DE = 1$ is circa 60% in both data sets. The binomial logistic regression analysis, specifically method ENTER was used for creation of the model.

Subsequent predictors were included in the analysis:

1. *Profitmaker/Lossmaker* (a dichotomous variable showing the final account balance of a trader);
2. *P/L* (a variable presenting profit or loss of the particular trade)
3. *Units* (a variable representing trades' volume);
4. *Length of trade* (mostly the explained variable, presenting length of trades in days);
5. *Gender* (a dichotomous variable, only "male" and "female" possibility was taken into account);
6. *Question No. 3* of the questionnaire (Have you ever had a feeling that you became "hooked-up" (addicted) to the game?);
7. *Question No. 5* of the questionnaire (Do you think that you are a person that enjoys taking risks?);
8. *Question No. 7* of the questionnaire (Have you ever made money by speculation?);
9. *Question No. 9* of the questionnaire (Have you ever done any adrenaline sports?).

Firstly the Omnibus Test of Model Coefficients, so-called "a goodness of fit test" was performed. This test provides an indication of how well the model performs over the one with none of predictors entered into the model. The full model containing all predictors was found statistically significant, as

$\chi^2(11, N = 8\ 693) = 748.85, p < 0.01$, indicating that the model was able to distinguish between traders who showed and did not show the disposition effect. The Hosmer–Lemeshow goodness of fit test is a poor fit test, indicated by the significance value smaller than 0.05. Also this test was found statistically significant, $Sig. < 0.05$, $\chi^2 = 236.24$ with 8 degrees of freedom. The Cox & Snell R Square and the Nagelkerke Square provide information about the amount of variation in the dependent variable explained by the model. In this model, the Cox $\&$ Snell R Square resulted in the height of 0.083 and Nagelkerke R Square in the height of 0.111, suggesting that between 8.3$\%$ and 11.1$\%$ of variability is explained by this model. The percentage of correctly predicted values is 66%, whereas the model predicted correctly 41$\%$ of "no disposition effect" values, and 83$\%$ of "disposition effect" values. The area under the curve (AUC) is equal to 0.7, which is according to Hosmer; Lemeshow; Sturdivant (2013) within the interval of acceptable discrimination, defined as $\langle 0.7, 0.8)$. Respective Gini coefficient is equal to 0.4. Several different updates were examined, such as different methods of data splitting and adjustments of the cut value, to increase the predicative ability of the model, however, none of them helped to achieve higher AUC and Gini coefficient.

Almost all variables were found statistically significant, except from the variable *Units*. Moreover, variable *P/L* has no impact on the explained variable. Hence the logistic regression model is defined as:

$$\ln\left(\frac{\text{p}(DE=1)}{\text{1-p}(DE=1)}\right) = 1.30 + 0.39 \cdot (\textit{profitmaker/loss maker = 1}) + 0.17 \cdot \textit{length of trade} - 0.97 \cdot Gender + \\ -0.66 \cdot (Q3 = 0.5) + 0.42 \cdot (Q3 = 1) - 0.43 \cdot (Q5 = 0.5) + 0.26 \cdot (Q5 = 1) - 0.48 \cdot (Q7 = 1) + 0.14 \cdot (Q9 = 1). \tag{5.1}$$

The strongest predictor was the answer "yes" to the *Question No.3*, focused on the addiction to trading. This answer increases the probability to achieve the $DE \geq 1.1$ by 52%. As described in Subchapter 4.1 of the thesis, 56% of students answered on that question that they did not feel addiction during the trading game. Looking at this result, we can see a close relationship of two irrational behaviors. A trader feeling addiction to trading is very close to gambling instead of making rational decisions, which deepens the tendency to go under the disposition effect, another irrational behavioral bias. The problematics of relationship between trading and gambling was studied[5] and discussed in Subchapter 2.3 of the thesis. Contrary, answering "sometimes" to this question leads to

[5] Cox et al. (2020), Pelster et al. (2018), or Konstantaras et al. (2011)

decrease of the disposition effect probability.

The second strongest predictor was variable *Profit maker/loss maker* showing an interesting result. According to this variable, being a profitmaker at the end of the trading period increases the probability of the disposition effect by almost 50%.

The variable connected with approach to speculation, *Question No.7*, showed that in case of no experience with speculative behavior the probability of showing the disposition effect is smaller. This leads to presumption that students having a previous tendency to speculate might be willing to accept the risk of waiting in case of loosing trades.

Regarding the *Gender*, it showed that males almost 60% times more likely showed the ratio *TimeLtoP* higher than 1.1, thus showed stronger tendency to the disposition effect. This result builds on the results of Subchapter 4.1 of the thesis, where the difference in tendency to succumb to the disposition effect from the gender point of view was statistically tested. Moreover, this is in line with previous research made by, for example, Jacobsen et al. (2014), Halko et al. (2012), Dwyera et al. (2002), or Barber et al. (2001), who proved females to be more risk averse than males, as the disposition effect is closely related to the risk attitude. Similar effect has also variable *Question No.5*, traders who answered "yes" to the question regarding their risk attitude seems to be willing to take a risk of waiting and face even higher loss in case of a loosing trade.

Contrary, *Question No.9*, referring about the adrenaline activities had opposite effect. Answering "no" to this question indicated smaller probability of achieving *TimeLtoP* ratio over 1.1.

The last variable with the smallest impact from all variables was the *Length of trade*, whereas the higher length of trade, the higher probability of the disposition effect.

Variable indicating the profitability and unprofitability of particular trades was also significantly affecting the dependent variable *DE* which can be interpreted as confirmation of increasing disposition effect by loss trade and loss aversion of traders, described by Kahneman et al. (1979).

Variables *Units* and *P/L* did not show any impact on the model.

To create the model only 80% of observations were used. The rest of data was used for validation of the model, which was done by substitution into the equation 5.1 and calculation of the probability of DE equal to 1. The AUC of this test sample resulted in the height of 0.43. According to this result, there is high probability that a negative case ($DE = 0$) will be marked as a positive case ($DE = 1$). This result can be caused by heterogeneity of the data, the impact of particular students was high. This result leads to question if other methods, for example a Gamma regression, would hep to achieve better results.

5.4 Conclusion

The thesis aimed to examine whether the disposition effect can be identified in selected experimental trading data and, if so, what were the affecting factors, and the differences within various groups of traders. The disposition effect is a crucial behavioral bias which needs to be understood to increase the possibility of defeating it. The goal was achieved using several methods, firstly the essential statistical testing of established hypotheses was used, followed by the application of the GUHA method and logistic regression analysis.

The thesis is of standard structure, starting with the Introduction, followed by Chapter 2, called Theoretical Framework of Research. As the name suggests, within this chapter, the theoretical framework of research was introduced. In this part, the reader could learn about the neoclassical approach to investment decision making, behavioral finance, and its key concept, the prospect theory proposed by Kahneman et al. (1979), where the disposition effect was outlined the very first time. Moreover, other studies that focused on the disposition effect[6], were described within this part of the thesis. However, most of the works used for examining the disposition effect data of real traders. More emphasis was placed on Weber et al. (1998), whose research based on experimental data collected from students. Finally, the rules of financial markets and exchange rates theory were described. The chapter was concluded by description of the OANDA FX Trading platform which was used for collecting the experimental trading data.

Subsequent Chapter 3 provides knowledge about the methodology used to achieve the research aim already mentioned hereabove. The particular methods, besides general statistical tests, were the GUHA method and logistic regression analysis. Moreover, the chapter also introduces the *TimeLtoP* ratio, created based on the duration approach to the disposition effect examination, which was used

[6]For example, Shefrin et al. (1985), Shiller et al. (1988), Shefrin et al. (1985), Odean (1998), Locke et al. (2005), Barberis et al. (2009), Cheng et al. (2013), Eom (2018), or Muhl et al. (2018).

in the empirical part of the thesis to assess the disposition effect. The chapter was concluded by the description of the data collection procedure.

Chapter 4 was focused on applying the theoretical knowledge and methodology, summarised in previous chapters. First of all, an underlying analysis of experimental data and statistical description was provided. Based on the results, the data does not follow normal distribution, neither after relevant adjustments as trimming nor log transformation. Nevertheless, it was still possible to use parametrical tests based on the Law of Large Numbers due to high number of observations.

First of all, four statistical hypotheses were tested. The author have focused on the general presence of the disposition effect in the experimental data along with the relation of the disposition effect to gender, trade size, and trader's home country. According to relevant statistical testing, the disposition effect was confirmed within examined data based on the length of trade and *TimeLtoP* examination and the impact of other tested characteristics. It was important to see, whether students succumb to the disposition effect in the same way real traders do. Even though they did not trade with real money, thus not facing a potential loss of their own funds, they appeared to be seriously loss averse. This fact leads to the presumption that disposition effect is an irrational behavioral bias, which does not depend on a real loss (trading on a demo account does not bear any real losses nor profits) but it is a part peoples' mindset.

Looking at the disposition effect and gender, the data confirmed work of several authors[7] thus it was possible to conclude that also students not trading with real money showed a statistically significant difference between genders (only two genders were taken into account). Firstly, according to the answers to questions, males more often answered that they enjoy taking risks as well as they have experienced an adrenaline sport. Secondly, looking at the *TimeLtoP* ratio used for quantifying the potential disposition effect, females showed more than twice as high values as males. This means that males held unprofitable trades, on average, 3.6x longer than profitable trades, but females held unprofitable trades 9.5x longer than profitable profitable ones. Moreover, females were trading on average with smaller amount; their trades were of smaller volumes than males' trades. The same result was also confirmed by Barber et al. (2001) and Cueva et al. (2019).

The holding period of trades concerning the trade volume also demonstrated statistically signif-

[7]For example, Halko et al. (2012), Dwyera et al. (2002), Barber et al. (2001), or Cheng et al. (2013).

icant differences. Small trades below the average volume, thus showed higher *TimeLtoP* than large trades, i.e. those over average. Considering the study of Eom (2018), who found empirical evidence for the opposite of the disposition effect and proposed that with the increasing absolute value of gains and losses the gap between holding period was decreasing, it may be concluded that, in this case, the same pattern was found.

Does home country matter? Yes, it does. The last hypothesis focused geographically, on the traders' country of origin. The background of this hypothesis assumed cultural differences causing differences in trading, as proposed by Wang et al. (2018) or Tan et al. (2019). As most of the traders were from New Zealand, the hypothesis was examining whether originating from New Zealand would have an effect on the tendency to succumb to the disposition effect. A statistically significant difference between students from New Zealand and other countries was found. The values were surprisingly high, as their holding period of loss trades was almost 50% higher when compared to students from other countries.

Subsequently, new hypotheses were generated by GUHA, which was done on original data and aggregated data. Trades were divided into three groups according to their length, and the main goal was to receive a hypothesis increasing the probability of belonging to a relevant interval. From all hypotheses generated by GUHA, only those with the highest *Lift* were chosen. Based on the results achieved from non-aggregated data, a loss making male, who did not feel addiction while trading, enjoys risks and has ever done adrenaline sports has more often shorter trade than others. Generally, this kind of behavior might be classified as speculation (see Tirole (1982) or Harrison et al. (1978)). Concerning hypotheses generated from aggregated data, GUHA provided several hypotheses increasing the probability of trading within the middle interval B2. However, any statistically significant hypotheses regarding the other intervals were received.

The last part of the empirical Chapter focused on examining the impact of particular variables on the disposition effect using the logistic regression analysis. The dependent variable *DE* equals to 1 if the *TimeLtoP* is higher than 1.1. Hence the disposition effect is assumed. Otherwise, the *DE* is equal to 0. The final model was not of the best characteristics, however, according to Hosmer; Lemeshow; Sturdivant (2013), it was still within the interval of acceptable discrimination. The outcome was aligned with the initial hypothesis testing. Logistic regression analysis confirmed the significance of addiction patterns on the disposition effect, i.e. one problematic, irrational behavior increased probability of another one. Moreover, the positive attitude to risk led to a higher probability of the disposition effect. Contrary to previous research and previous statistical testing within this thesis, the model demonstrates that males have higher probability of succumbing to the disposition effect than females.

Overall, the disposition effect has been confirmed within the experimental data set and examined from several perspectives, thus the aim of the thesis was fulfilled. Given the results, it may be assumed that traders, no matter if they face real losses or not, succumb to the same behavioral bias, while the degree of this behavioral bias effect is mainly affected by gender, risk attitude, and culture. The research also opened several potential questions for further research and examination, the main future focus should be on the passively closed trades and its comparison to behavioral patterns found within actively trades, which might lead also to research focused on AI trading and potential behavioral patterns coded in algorithms using for it.

6 References

ADAMS, Renée B.; FUNK, Patricia, 2012. Beyond the Glass Ceiling: Does Gender Matter? *Management Science*. **58**(2), 219–235.

BARBER, Brad M.; ODEAN, Terrance, 2001. Boys will be boys: Gender, overconfidence, and common stock investment. *The Quarterly Journal of Economics*. **116**(55), 261–292.

BARBERIS, Nicolas; XIONG, Wei, 2009. What Drives the Disposition Effect? An Analysis of a Long-Standing Preference-Based Explanation. *The Journal of Finance*. **64**(2), 751–784.

CHEN, Gong-meng; KIM, Kenneth A.; NOSFINGER, John R.; RUI, Oliver M., 2007. Trading performance, disposition effect, overconfidence, representativeness bias, and experience of emerging market investors. *Journal of Behavioral Decision Making*. **20**(4), 425–451.

CHENG, Teng Y.; LEE, Chun I.; LIN, Chao H., 2013. An examination of the relationship between the disposition effect and gender, age, the traded security, and bull – bear market conditions. *Journal of Empirical Finance*. **21**, 195–213.

COX, Ruben; KAMOLSAREERATANA, Atcha; KOUWENBERG, Roy, 2020. Compulsive gambling in the financial markets: Evidence from two investor surveys. *Journal of Banking & Finance*. **111**.

CUEVA, Carlos; ITURBE-ORMAETXE, Inigo; PONTI, Giovanni; TOMAS, Josefa, 2019. Boys will still be boys: Gender differences in trading activity are not due to differences in (over)confidence. *Journal of Economic Behavior Organization*. **160**, 100–120.

DVORACKOVA, Hana; JOCHEC, Marek, 2018. Disposition Effect in Currency Trading: An Evidence from Experimental Student Games. In: *9th International Scientific Conference Managing and Modelling of Financial Risks*. Ostrava: VŠB–Technical University Ostrava.

DWYERA, Peggy D.; GILKESON, James H.; LIST, John A., 2002. Gender differences in revealed risk taking: evidence from mutual fund investors. *Economics Letters*. **76**(2), 151–158.

EOM, Yunsung, 2018. The opposite disposition effect: Evidence from the Korean stock index futures market. *Finance Research Letters*. **26**, 261–265.

FIELD, Andy P., 2013. *Discovering statistics using IBM SPSS Statistics: and sex and drugs and rock 'n' roll*. 4th ed. London: SAGE.

GAINSBURY, Sally M.; KING, Daniel L.; RUSSELL, Alex M.T.; DELFABBRO, Paul; HING, Nerilee, 2017. Virtual addictions: An examination of problematic social casino game use among at-risk gamblers. *Addictive Behaviors*. **64**, 334–339.

GOLDREICH, Oded, 2010. *P, NP, and NP-Completeness: The Basics of Complexity Theory*. New York: Cambridge University Press.

HÁJEK, Petr, 2003. Relations in GUHA Style Data Mining. In: *7th International Seminar on Relational Methods in Computer Science*. Bad Malente: RelMiCS.

HÁJEK, Petr; HAVEL, Ivan; CHYTIL, Milan, 1966. The GUHA method of Automatic Hypotheses Determination. *Computing 1*. **4**, 293–308.

HÁJEK, Petr; HOLEŇA, Martin; RAUCH, Jan, 2010. The GUHA method and its meaning for data mining. *Journal of Computer and System Sciences*. **76**(1), 34–48.

HALKO, Marja-Liisa; KAUSTIA, Markku; ALANKO, Elias, 2012. The gender effect in risky asset holding. *Journal of Economic Behavior & Organisation*. **83**(1), 66–81.

HARRISON, J. Michael; KREPS, David M., 1978. Speculative Investor Behavior in a Stock Market with Heterogeneous Expectations. *Quarterly Journal of Economics*. **92**(2), 323–336.

HOSMER, David W.; LEMESHOW, Stanley, 2004. *Applied Logistic Regression*. New Jersey: John Wiley & Sons.

HOSMER, David W.; LEMESHOW, Stanley; STURDIVANT, Rodney X., 2013. *Applied Logistic Regression*. 3rd ed. New Jersey: John Wiley & Sons.

JACOBSEN, Ben; LEE, John B.; MARQUERING, Wessel; ZHANG, Cherry Y., 2014. Gender differences in optimism and asset allocation. *Journal of Economic Behavior & Organisation*. **107**, 630–651.

KAHNEMAN, Daniel; TVERSKY, Amos, 1979. Prospect theory: An analysis of decision under risk. *Econometrica*. **46**, 171–185.

KAUSTIA, Markku, 2013. *Behavioral Finance: Investors, Corporations, and Markets*. 3rd ed. New Jersey: John Wiley & Sons.

KONSTANTARAS, Konstantinos; PIPEROPOULOU, Natasha, 2011. Compulsive gambling and the underestimation of risk. *European Psychiatry*. **26**, 324–336.

LEHENKARI, Mirjam, 2012. In Search of the Underlying Mechanism of the Disposition Effect. *Journal of Behavioral Decision Making*. **25**(2), 196–209.

LISp-Miner [online]. Available also from: `http://lispminer.vse.cz/index.html`.

LOCKE, Peter R.; MANN, Steven, 2005. Professional trader discipline and trade disposition. *Journal of Financial Economics*. **76**(2), 401–444.

MANCINI, Loriano; RANALDO, Angelo; WRAMPELMEYER, Jan, 2013. Liquidity in the foreign exchange market: measurement, commonality, and risk premiums. *Journal of Finance*. **68**(5), 1805–1841.

MARCIUKAITYTE, Dalia; SZEWCZYK, Samuel H., 2012. Disposition Effect and Public Secondary Equity Offerings. *Journal of Financial Research*. **35**(1), 55–77.

MUHL, Stefan; TALPSEPP, Tõnn, 2018. Faster learning in troubled times: How market conditions affect the disposition effect. *The Quarterly Review of Economics and Finance*. **68**, 226–236.

ODEAN, Terrance, 1998. Are investors reluctant to realize their losses. *Journal of Finance*. **53**(5), 1775–1798.

PELSTER, Matthias; HOFMANN, Annette, 2018. About the fear of reputational loss: Social trading and the disposition effect. *Journal of Banking and Finance*. **94**, 75–88.

PUDLÁK, Pavel; SPRINGSTEEL, Frederic N., 1979. Complexity in Mechanized Hypothesis Formation. *Theoretical Computer Science*. **8**(2), 203–225.

SARMIENTO, Julio; RENDÓN, Jairo; SANDOVAL, Juan S.; CAYON, Edgardo, 2019. The disposition effect and the relevance of the reference period: Evidence among sophisticated investors. *Journal of Behavioral and Experimental Finance*. **24**, 100–111.

SHEFRIN, Hersh; STATMAN, Meir, 1985. The disposition to sell winners too early and ride losers too long: Theory and evidence. *Journal of Finance*. **40**(3), 777–790.

SHILLER, Robert; CASE, Karl, 1988. *The Behavior of Home Buyers in Boom and Post-Boom Markets*. Cowles Foundation Discussion Papers.

STANLEY, Morgan, 2012. Social gambling: click here to play. *Morgan Stanley Research: Global*. **23**, 171–175.

TAN, Gary; CHEONG, Chee Seng; ZURBRUEGG, Ralf, 2019. National culture and individual trading behavior. *Journal of Banking Finance*. **106**, 357–370.

TIROLE, Jean, 1982. On the Possibility of Speculation under Rational Expectations. *Econometrica*. **50**(5), 1163–1181.

TURUNEN, Esko, 2009. Interpreting GUHA data mining logic in paraconsistent fuzzy logic framework. *Lecture Notes in Computer Science*. **5783**, 284–293.

WANG, Jianxin; HOUSER, Daniel; XU, Hui, 2018. Culture, gender and asset prices: Experimental evidence from the U.S. and China. *Journal of Economic Behavior Organization*. **155**, 253–287.

WEBER, Martin; CAMERER, Colin F., 1998. The disposition effect in securities trading: an experimental analysis. *Journal of Economic Behavior & Organisation*. **33**, 167–184.

7 List of Author's Publications

Papers

DVOŘÁČKOVÁ, H., M. JOCHEC and T. TICHÝ. Disposition Effect in Currency Trading: An Evidence from Experimental Student Games. *Review of Research and Social Intervention*, 2019, 64, pp. 246-261. ISSN 1583-3410.

DVOŘÁČKOVÁ, H., A. KRESTA and T. TICHÝ. Sales prediction in Ice category with an LFL-Forecaster applied to quarterly temperatures. *Ekonomická revue – Central European Review of Economic Issues*, 2018, 21 (2), pp. 33–39. ISSN 1212-3951.

Conference Proceedings

DVOŘÁČKOVÁ, Hana. Experimental finance: the Gender Differences in the Disposition Effect Bias. In: *Mathematical Methods in Economics 2019: 37th International Conference*. České Budějovice: University of South Bohemia in České Budějovice, 2019. ISBN 978-80-7394-760-6.

DVOŘÁČKOVÁ, Hana and Marek JOCHEC. Examination of the Gender Differences in the Experimental Trading under the Association Analysis Method. In: *18th Conference on Information and Management Science* , Lhasa, China, 2019.

DVOŘÁČKOVÁ, H., M. JOCHEC and T. TICHÝ. Disposition Effect in Currency Trading: An Evidence from Experimental Student Games. In: *12th International Conference of Iranian Operational Research.* Babolsar: Iranian Operational Research Society, 2019.

DVOŘÁČKOVÁ, H., M. JOCHEC and T. TICHÝ. Disposition Effect in FX Trading: Empirical from Experimental Student Games. In: *Finance Research Symposium 2019*. Nur-sultan: Nazarbayev University Nur-sultan, Kazakhstan, 2019.

DVOŘÁČKOVÁ, Hana and Marek JOCHEC. Different Points of View on the Disposition Effect in the Experimental Currency Trading. In: *MEKON 2019: 21st International Conference.* Ostrava: VSB – TUO, 2019.

DVOŘÁČKOVÁ, Hana and Marek JOCHEC. Disposition Effect in the Experimental Currency Trading. In: *9th International Scientific Conference Managing and Modelling of Financial Risks*. Ostrava: VSB – TUO, 2018. ISBN 978-80-248-4225-7.

DVOŘÁČKOVÁ, Hana and Marek JOCHEC. The Gender Differences in the FX Trading Approach. In: *MEKON 2018: 20th International Conference*. Ostrava: VSB – Technical Univesity of Ostrava, 2018.

DVOŘÁČKOVÁ, Hana and Marek JOCHEC. Evaluation of the Behavioral Differences in the FX Trading Approach with Regard to the Gender. In: *11th International Scientific Conference on Financial Management of Firms and Financial Institutuions*. Ostrava: VSB – TUO, 2017. ISBN 978-80-248-4138-0.

DVOŘÁČKOVÁ, Hana. Behavioral Finance: The Gender Differences in the FX Trading Approach. In: *ECEE 2017: Economic Challengees in Enlarged Europe*. Tallin: Tallin University of Technology, 2017. ISSN 2382-6797.

DVOŘÁČKOVÁ, Hana. Analysis of Impacts on the Profit Making on the FX Market Using the Logistic Regression Model. In: *MEKON 2017: 19th International Conference*. Ostrava: VSB – TUO, 2017.

DVOŘÁČKOVÁ Hana, and Marek JOCHEC. Evaluation of the Impact of Selected Variables on the Investment Behavior. In: *Managing and Modelling of Financial Risks: 8th International Scientific Conference.* Ostrava: VSB – TUO, 2016. ISBN 978-80-248-3994-3.

8 Summary

The doctoral thesis focuses on behavioral finance, specifically on the disposition effect in trading and its presence in an experimental data set. The disposition effect was firstly described in the stock-investing context as a behavioral tendency of investors to hold on losing stocks for too long and sell winning stocks too early. Simply put, an investor in the case of a loss trade hopes that the situation reverses in the near future. Contrary in case of profitable trades, the investor is afraid of a potential loss that might come if he holds the trade longer. It was mentioned for the first time in work called *Prospect Theory: An Analysis of Decision under Risk*, published by Kahneman et al. (1979).

The disposition effect is a significant phenomenon, which might lead in the worst case to a global financial market disorder. The thesis aims to examine whether the disposition effect can be identified in selected experimental trading data and, if so, what are the affecting factors and the differences within various groups of traders.

The use of experimental data and the procedure if their completion are significant differences of the thesis as compared to previous works, which, to author's best knowledge, were based mostly on data collected from real investors, for which specific characteristics and standardized behavioral biases such as self-selection bias (see Adams et al. (2012)) can be assumed. Moreover, the effect of the trader's current wealth or social status is eliminated. Additional advantages and disadvantages of the experimental setting of the research are presented in Chapter 3.5 of the doctoral thesis.

Overall, the disposition effect has been confirmed within this data set and discussed from several points of view. The key findings to be that traders, no matter if they face real losses or not, succumb to the disposition effect. Moreover, the disposition effect is closely related and affected by the risk attitude, gender, and culture.

The doctoral thesis has the standard structure. The Introduction is followed by Chapter 2, proposing a theoretical background of the research and review of relevant literature. Chapter 3 focuses on mathematical and statistical methods used to achieve the aim, namely statistical methods, GUHA method, and logistic regression. Moreover, the *TimeLtoP* ratio and the process of experimental data collection are described. A crucial part of the thesis, Chapter 4, presents the results of the research. The thesis is, as usual, finished by a Conclusion.

key words: behavioral finance, experimental finance, disposition effect, GUHA, gender differences, cultural differences