Report on Master's Thesis

Institute of Economic Studies, Faculty of Social Sciences, Charles University

Student:	**Bc. Tereza Javůrková**
Advisor:	**PhDr. Jiří Kukačka, Ph.D.**
Title of the thesis:	**Gambler's Fallacy in Investor's Decision-making**

OVERALL ASSESSMENT *(provided in English, Czech, or Slovak):*

Short summary

This thesis explores the impact of the Gambler's Fallacy on investor behavior in the stock market and aims to integrate psychological research on this behavioral phenomenon into the field of finance. The objective is to analyze how human misconceptions about the probabilities of independent events influence the dynamics of the stock market. Specifically, the study examines the profitability of two types of virtual investors whose decision-making is influenced by distorted probabilities derived from the Gambler's Fallacy. Additionally, two other benchmark investors' strategies with varying levels of randomness are defined for comparison. The analysis investigates the gains of investors in both a simulated efficient market and the real S&P 500 index constituents. The research methodology employs three approaches: simulation analysis, empirical frequency analysis, and asset pricing models.

Through the simulation approach and frequency analysis applied to historical stock prices, it is discovered that investors affected by the Gambler's Fallacy achieve statistically higher returns compared to random investors. Furthermore, the study employs both the three-factor and five-factor Fama & French asset pricing models on stocks grouped into portfolios based on their previous earnings per share trends. The results reveal a negative excess return for stocks that, based on recent trends, are more likely to be purchased by investors exhibiting a bias towards the Gambler's Fallacy. These findings align with a novel asset pricing approach derived directly from psychological research.

Contribution

The thesis aims to investigate asset pricing models by considering the influence of psychological phenomena related to human misconceptions, specifically the tendency to search for long-term effects even in short sequences. It presents the concept of the Law of Small Numbers, which suggests that individuals may exhibit biased decision-making due to their inability to understand the nature of independent events in the short run. The author focuses on identifying this phenomenon among stock market agents and proposes four types of investors, two of which are specifically influenced by the Gambler's Fallacy, a concept derived from the Law of Small Numbers. The main contribution of the thesis lies in conducting simulations using real data, including frequency analysis and the application of the three-factor and five-factor Fama and French models.

Although the thesis presents intriguing insights that are likely to generate interest in the topic, I have several strong objections to the materials presented.

Chapter 4 of the thesis provides an analysis of S&P500 data, which I find to be simplistic. The methods used in the analysis overlook important factors known to influence changes in stock prices, such as company performance, industry performance, global events and macroeconomic factors, market sentiment, company news and events, and government policies and regulations. However, it appears that the author acknowledges this limitation in a later comment in Section 5.

The frequency analysis presented in the thesis focuses on sequences of stock price increases and declines, but it is conducted on a finite dataset spanning only 10 years. The author draws strong conclusions regarding the efficient market hypothesis, but fails to acknowledge that the results may still be consistent with it, considering the short time series where the effects of the Law of Large Numbers may not be sufficiently evident. It would have been appropriate for the student to perform a

Report on Master's Thesis

Institute of Economic Studies, Faculty of Social Sciences, Charles University

Student:	**Bc. Tereza Javůrková**
Advisor:	**PhDr. Jiří Kukačka, Ph.D.**
Title of the thesis:	**Gambler's Fallacy in Investor's Decision-making**

Chi-squared Test of Goodness of Fit to determine whether the data provide evidence to reject the hypothesis of equal frequencies.

The analysis conducted in Chapter 5 appears to be disjointed from the calculations performed in Chapter 4. The time series considered in Chapter 5 spans 20 years, while the models utilize data with different frequencies (weekly data in Chapter 4 and quarterly data in Chapter 5). Additionally, the models in Chapter 5 incorporate the factors mentioned earlier to a significant extent. As a result, any attempt to compare the results becomes incompatible due to these differences.

Moreover, in Chapter 5, the author constructs various monthly time series, such as an average abnormal return of the spread, a time series of average returns of the spread, and a monthly time series consisting of the difference between the average returns of stocks. In Section 5.3, the author claims to transform the data set to a quarterly time series by selecting only data from March, June, September, and December. However, it remains unclear what motivates this transformation and how it affects the interpretation of the achieved results.

Furthermore, the confusion regarding the content of Chapter 5 is compounded by the author's statement of "eliminating 46 stocks" and arriving at a "final data set containing 455 S&P 500 index stocks" (46 + 455 = 501?). This inconsistency undermines the contribution of this novel approach.

Overall, I find it unclear why the author chose to analyze the raw data set without removing trends or other systemic components before attempting to identify a random walk in the residual series. Additionally, I do not understand the rationale behind using different data sets for the numerical experiments. Consequently, it comes as no surprise that the author herself appears puzzled by the summary of the results in Section 5.4.1.

Methods

Although the chosen methods in the thesis are typical for a Master's level at IES, I believe there are noticeable shortcomings that should have been addressed in the final manuscript. As previously mentioned, the numerical analysis in Chapter 4 could have been enhanced by including a Goodness of Fit analysis. Additionally, the regression analysis in Chapter 5 lacks a thorough discussion of the underlying assumptions necessary to ensure the reliability of the calculated estimates.

Furthermore, as expressed in the previous section, I am dissatisfied with how the data is treated in both Chapters 4 and 5.

Literature

The literature review is rather brief and on occasions does not cite the appropriate key literature sources. Although the book by Andel (2002) is certainly a highly appraised textbook on (basic) mathematical statistics, I would expect (by more than 100 years) older sources for Law of Large Numbers or Central Limit Theorem, e.g. works of Pafnuty Chebyshev, Émile Borel, Andrei Kolmogorov, Pierre-Simon Laplace or Aleksandr Lyapunov. Using a 2017 paper published in Korean Journal of Anestesiology as a source for Central Limit Theorem seems disrespectful.

Manuscript form

With the exception of a somewhat abrupt transition between Chapters 4 and 5, the manuscript demonstrates a logical structure. The thesis is written in reasonably good English, although there are

Report on Master's Thesis

Institute of Economic Studies, Faculty of Social Sciences, Charles University

Student:	**Bc. Tereza Javůrková**
Advisor:	**PhDr. Jiří Kukačka, Ph.D.**
Title of the thesis:	**Gambler's Fallacy in Investor's Decision-making**

typos that occur with noticable frequency. Unfortunately, these typos sometimes affect important terminology, as seen with the mention of the "Week form" of the Efficient Market Hypothesis on page 8. However, the list of literature is properly formatted. Additionally, the thesis incorporates high-quality graphics and presents tables that are easy to read, which I appreciate.

Overall evaluation and suggested questions for the discussion during the defense

The results of the Turnitin analysis do not indicate significant text similarity with other available sources.

Taking into account the arguments above, in my view, the thesis fulfills the requirements for a bachelor thesis at IES, Faculty of Social Sciences, Charles University, I recommend it for the defense and suggest a grade A.

I propose the following to be discussed during the thesis defense:

1/ Why did you choose to perform the frequential analysis on data covering 2012-2022 and not 2002-2022 as in analysis in Chapter 5?

2/ Except for the differences in length of the data sets, how would you modify the frequential analysis performed in Chapter 4 in order to achieve (higher) comparability of the outcomes of the two considered approaches?

3/ What other psychological phenomena, along with Gambler's Fallacy, could be studied (after corresponding modifications) using your proposed method?

SUMMARY OF POINTS AWARDED *(for details, see below):*

CATEGORY		**POINTS**
Contribution	*(max. 30 points)*	22
Methods	*(max. 30 points)*	18
Literature	*(max. 20 points)*	17
Manuscript Form	*(max. 20 points)*	18
TOTAL POINTS	*(max. 100 points)*	**75**
GRADE	**(A – B – C – D – E – F)**	**C**

NAME OF THE REFEREE: Michal Červinka

DATE OF EVALUATION: June 19, 2023

Michal Červinka

Digitally signed by Michal Červinka
Date: 2023.06.19 14:59:00 +02'00'

Referee Signature

EXPLANATION OF CATEGORIES AND SCALE:

CONTRIBUTION: *The author presents original ideas on the topic demonstrating critical thinking and ability to draw conclusions based on the knowledge of relevant theory and empirics. There is a distinct value added of the thesis.*

METHODS: *The tools used are relevant to the research question being investigated, and adequate to the author's level of studies. The thesis topic is comprehensively analyzed.*

LITERATURE REVIEW: *The thesis demonstrates author's full understanding and command of recent literature. The author quotes relevant literature in a proper way.*

MANUSCRIPT FORM: *The thesis is well structured. The student uses appropriate language and style, including academic format for graphs and tables. The text effectively refers to graphs and tables and disposes with a complete bibliography.*

Overall grading:

TOTAL	**GRADE**
91 – 100	**A**
81 - 90	**B**
71 - 80	**C**
61 – 70	**D**
51 – 60	**E**
0 – 50	**F**