VILNIUS UNIVERSITY

FACULTY OF ECONOMICS AND BUSINESS ADMINISTRATION

KRISTIJONAS VĖLYVIS

Quantitative Economics

Bachelor Thesis

EUROLEAGUE BETTING MARKET AND ITS EFFICIENCY

Supervisor Milda Norkutė

Date of submission of the academic paper 2021 05 10

Registration No. 141868

Assessment of the academic paper

Vilnius, 2021 05 10

Table of Contents

1. INTRODUCTION

The betting markets have grown rapidly in the last few decades. However, the first appearances of gambling reach as far back as 500 BC when a game of dice was invented. Later, there used to be various restrictions and regulations for gambling, until the 20th century when it became more mainstream and popular, as an increasing number of casinos began to open. Nevertheless, this paper concentrates on online betting, which began to exist in the 1990s and has grown exponentially since the beginning of this millennium. One of the principal reasons for its thriving is rapid technology improvement as bettors could transfer their hobby to the internet, which resulted in a 10% yearly growth of Europe's online betting market (European Gaming and Betting Association, 2021). As of 2018, the EU records almost half of the global online gambling market. Also, the US market is gaining momentum as well due to recent legalizations in some states, such as Nevada, Pennsylvania, or New Jersey, which is the largest market in the US. Thus, the legal exemptions are another huge factor for the growth of betting markets. Additionally, online betting has experienced a significant upswing in the Baltics, as, for instance, a 47.1% boom in Lithuania's online gambling revenue was recorded in 2020 (Narayan, 2021). The rise of online betting resulted in an enlarged number of firms that deliver gambling services (i.e., bookmakers), which, consequently, contributed to at first a small but gradually increasing amount of academic literature on this topic (Radiul & Janusonis, 2020).

There are a few important aspects to why betting markets have been highly researched. First, gambling markets might be compared to typical financial markets, and, therefore, can be useful in analyzing market efficiency (Vlastakis, Dotsis & Markellos, 2009). Wagering (i.e., betting) markets, however, have some advantageous differences over the financial markets; for example, a bet placed by a punter "has a well-defined termination point at which its value becomes certain", which makes it a lot easier to test for rationality and efficiency in the market (Thaler & Ziemba, 1988). The other reason for the increased amount of analysis of a betting market is its questionable efficiency itself, as researchers find that bookmakers, by forming odds (i.e., measurements of likelihood of various outcomes), might create some opportunities for consistent profitable betting strategies which, theoretically, shall not exist (Gray & Gray, 1997; Vlastakis et al., 2009).

In theory, the strictest argument for efficiency of a betting market is that the market prices (i.e., the odds set by a bookmaker) are the most precise forecasts of an event having all the available information incorporated in it (Sauer, 1998; Radiul & Janusonis, 2020). This means that neither bookmaker nor punter might earn abnormal profits that are higher than the bookmaker's profit margin (i.e., commission tax). In other words, there should not exist a profitable betting strategy. However, the majority of academic papers on betting markets have delivered contrary conclusions. A transaction in betting market consists of two parties, one being a bookmaker, and the other – a punter. Most of the analyses concentrate on the existence of various biases, such as home-away or favorite-longshot, which both parties are subject to. This manifests by overestimating some outcomes of events and underestimating the others, and, thus, creating mispricings, which is one of the key factors for market's inefficiency (Goddard & Asimakopoulos, 2004).

Therefore, this paper aims to reveal whether EuroLeague betting market is efficient. EuroLeague, which is the second-best basketball league in the world, was chosen because its betting market has not been investigated yet. Additionally, it contains teams from approximately 10 different countries in Europe, which, as already mentioned, has the highest share of online gambling markets worldwide. In this spirit, the first objective of this study is to identify which behavioral biases are present in the market. The existence of different biases in a gambling market is one of the key measurements of market's efficiency because, according to Gray & Gray (1997), if either bookmakers or bettors can take advantage of systematic biases (e.g., favorite-longshot bias) to yield consistent returns, it is an evidence of an inefficient market. Furthermore, the second objective to test for efficiency of the market is to evaluate various betting strategies. Therefore, a betting simulation was conducted to reveal whether there exist any profitable wagering strategies in EuroLeague betting market. Again, for a gambling market to be considered efficient, punters cannot obtain profits higher than the bookmakers' over-round (i.e., profit margin).

To begin with, various econometric model specifications are considered to investigate if any observable variables are significant in predicting the outcomes of sports events. Almost every examined model presented a huge significance for home and favorites dummy variables, and, unsurprisingly, the main conclusion was drawn that the home-favorites (i.e., teams that are playing at home-field and have a higher probability to win) have the biggest advantage to win a game. Furthermore, seasonal winning percentage and average point difference were also significant in forecasting outcomes of future events, which revealed a weak-form inefficiency for EuroLeague betting market.

The second approach for proving the inefficiency of this market was by examining the most popular biases that appear in the academic literature about betting markets, such as home-away, favorite-longshot, and most-likely/least-likely. A betting simulation was conducted to compare potential profits for punters if they bet the exact same value on various outcomes. The results showed that all of the mentioned biases exist in the EuroLeague betting market, and an appearance of home-favorite bias was again proved to exist.

Thus, the structure of this paper is as follows: the first section, literature review, describes different types of betting, discusses gambling market efficiency and common biases, and presents how betting markets are tested for efficiency in the academic literature. In the second section, the data and methodology are presented that have been used for the econometric part of the paper. The third section discusses the main findings of the analysis, and, finally, the last section concludes.

2. LITERATURE REVIEW

2.1. Theoretical context of the betting markets

2.1.1. Mechanisms of fixed-odds wagering

Betting market refers to a so-called speculative market, where investors (i.e., punters) are involved in money trades that are risky but have potentially high returns (Franck, Verbeek & Nuesch, 2013). Additionally, wagerers are free to choose whether they are willing to be involved in the transactions, similarly as in quote-driven markets (Flepp, Nuesch & Franck, 2017, Radiul & Janusonis, 2020). These types of trading consist of contracts, where the cash flow depends on the outcome of a specific event, most commonly the outcome of a sports match. There are a few types of betting, such as fixed-odds, pari-mutuel, and odds spread wagering. In the fixed-odds case, the cash flow and the payout for the wagerers are determined by the odds (Franck et al., 2013).

For fixed-odds betting, there exist a few market mechanisms, namely the bet exchange market and the bookmaker. The latter is the most common type of mechanism for the sports betting markets (Franck et al., 2013). In this case, the bookmakers, who act as an intermediary, form the price (i.e., odds) for a sports match usually a few days in advance, and the wagerers might choose what to put their bets on (Radiul & Janusonis, 2020). Additionally, the odds may vary before an event; however, when a bet is placed, no modifications from either side can be adjusted.

There are a lot of discussions in the economic literature about the efficiency of stock markets. According to Thaler and Ziemba (1998), the structure of gambling markets allows them to be "better suited for testing market efficiency and rationality", as each contract (a bet) has an exact expiry date, after which its initial worth becomes certain. In stock markets, however, the intrinsic value of a property is determined not only by its present value and future cash flows but also by its future price, which is the main reason why testing the efficiency of stock markets become more difficult comparingly to the betting markets (Thaler & Ziemba; 1998, Radiul & Janusonis, 2020).

Furthermore, as said before, the other tool of the fixed-odds gambling markets is the exchange market. In this case, the market allows bettors to compete against each other, as anyone willing to participate in the transactions might not only buy contracts but sell them as well. Thus, the bookmakers in this type of mechanism do not represent the opposite party but rather just collect a commission fee for any kind of transaction made between the punters. (Radiul & Janusonis, 2020). However, this study concentrates on the bookmaker type of mechanism of the fixed-odds betting, where odds are not controlled by demand and supply of the wagerers (as in the latter example) but rather are set by the bookmakers (Franck et al., 2013).

2.1.2. Various forms of gambling

In theory, there can be various forms of sports gambling, such as pari-mutuel, odds spread wagering, or, as described in the section above, fixed-odds betting.

First, pari-mutuel betting form is most common in sports where the competitors are ranked, such as horse- or dog-racing competitions. In this case, there is a pool of money from bets on all the possible outcomes, which is later distributed to punters who placed their bets on the winner, proportionally (Ottaviani & Sorensen, 2008). Therefore, each wagerer represents either demand or supply side – demanding a portion of the winning cake in case of success and contributing money to the other punters when their choice was not profitable.

Second, odds spread wagering can be used in sports with a point system, such as basketball or football. In this case, the odds are formed in a way that they represent a probability that a team of record would beat the predetermined point spread, i.e., would win by a certain number of points (Teall, 2018). Therefore, a punter earns his payout when if, for example, his chosen team outscored the opponents by more points than the spread set by the bookmaker (Radiul & Janusonis, 2020). This method of wagering is convenient in some instances where a top team has a massive advantage against its opponent, as bookmakers might find it more profitable to set a point spread boundary at which various wagerers would be stimulated to place bets (Gray & Gray, 1997; Teall, 2018).

Lastly, fixed-odds gambling, which this study is oriented to, is the most popular form of sports betting (Franck et al., 2013). Initially, bets for every outcome of the game (home win and away win or draw in some sports) are preset by the bookmakers usually a few days before an event. They might fluctuate over time due to various circumstances, such as, for instance, changes in teams' rosters; however, once a wagerer places a bet, it remains valid, and the payout is paid to the bettor in case of success. Nevertheless, the fixed odds do not represent the true probabilities of an event since bookmakers are profit-maximizing institutions that seek for profit margin, also known as the "over-round".

2.2. Efficiency

Literature on efficiency of the gambling markets has been existing for more than 30 years. One of the initial researches on that was conducted by Thaler and Ziemba (1998), who argued that the betting markets have the potential to be more efficient. Theoretically, a gambling market is considered to be efficient is when the market prices – in this case, the odds set by a bookmaker – are the most accurate forecasts of an event having all the available information incorporated in it (Sauer, 1998; Radiul & Janusonis, 2020). This means that neither bookmaker nor wagerer might yield average payoffs higher than the commission tax; therefore, wagerers cannot sustain returns that contrast with the bookmaker's profit margin, while bookmakers cannot earn a higher margin than their competitors (Vlastakis et al., 2009). In general, testing for betting markets efficiency is much less complicated comparingly to financial markets due to the fact that market prices "have a well-defined period of life at the end of which their value becomes certain" (Vlastakis et al., 2009).

In financial markets, there are three different types of market's efficiency defined by Fama (1970): weak-form, semi-strong, and strong form efficiency. First, as Dowie (1976) suggests, a stock market can be regarded as efficient "if its prices always fully reflect available information". Available information, however, can be distributed into three groups – historical prices, other publicly available information (e.g., issues), and private information (all the possible information that is available to certain individuals only), and in turn they correspond to the three types of market efficiency, respectively (Dowie, 1976). Therefore, when testing for market's efficiency, it is considered to be weakly efficient when the information set consists of the historical prices only; semi-strongly efficiency occurs when the information set includes other public information; and strongly efficient market appears when any investors "have monopolistic access to any information relevant for prices formation" (Fama, 1970). In conclusion, any profits shall be lower than the transaction costs in a case of an efficient market.

The same theory might be applied to the betting markets. If a wagering market is considered to be weakly efficient, then neither bookmakers nor wagerers, operating with only historical archives of odds, cannot yield abnormal returns (Radiul & Janusonis, 2020). Accordingly, neither side of the transaction can obtain abnormal returns using any public information in a market that is semi-strong efficient, as the public information will not improve the ability to forecast outcomes of future events. The same theory applies to the strongly efficient betting markets, which is not apparent in practice, where neither public nor private information would not give an edge to any participant of the market (Radiul & Janusonis, 2020).

Finally, Gray and Gray (1997) discuss about other ways to test for betting market's efficiency. First, they analyze the appearance of statistical efficiency, which states that the preset odds for a sports match have to reflect the true probabilities, and, if any observable information is available to predict the outcomes more accurately, the market, therefore, is concluded to be statistically inefficient. Furthermore, they suggest an economic form of efficiency for the gambling markets. They argue that present behavioral biases are "in self, evidence of inefficiency" (Gray & Gray, 1997). Consequently, if any party can make use of systematic biases (e.g., home-away bias) to obtain consistent returns, it reveals an evidence of economically inefficient market.

2.3. Regression models

Furthermore, this analysis represents some econometric models with various variables to check whether any publicly available data is useful for predicting the outcomes of upcoming matches. There are numerous different attempts to test for betting market's efficiency in the literature. In this study, the betting market efficiency is assessed using probit and logit regressions with binary dependent variable, which is equal to one if the predetermined favorite team wins a game and zero otherwise. The use of a point spread as dependent variable is not possible, however, because the historical data of odds only are available publicly.

The initial augmented model, which consists of home and favorite dummies, was implemented by Gray and Gray (1997), where they used a point spread as a dependent variable. In that paper, they found that the home teams and longshots (teams that have less than 50% probability to win) have the advantage on beating the point spread against the away teams and favorites, which partially conform this research, where similar conclusions for the home team effect were drawn. As there was evidence of systematic biases existence, (e.g., favoritism of home teams), NFL betting market was considered to be inefficient because, according to Gray and Gray (1997), there was an appearance of biases in the market that could be useful in creating a profitable betting strategy. Nevertheless, the authors found that some of the biases, for example, a home-underdog bias, tended to disperse over time.

Additionally, a similar probit model was analyzed by Harkins (2013), where the underdog effect was found to be statistically significant, but the home field effect turned out to be the opposite, which, again, complied with my analysis only to a limited degree. In the end, Harkins (2013) shows that seasonal cumulative point differential was not significant in the NFL betting market, which suggested that bookmakers might take it into consideration while setting the odds, suggesting NFL gambling market to operate at a weak-form efficiency.

Moreover, WNBA (Women NBA) betting market was investigated by Paul and Weinbach (2014), who presented ambiguous results. Correspondingly to Gray and Gray (1997) model, they examined the prediction of game outcomes using point spread as a regressor and conducted a betting simulation similar to Constantinou and Fenton (2013) test, which is discussed in the next section. Later, they supplemented the initial model with road-favorite dummy, which was of high significance, contributing to Vlastakis et al. (2009) findings of away-favorite bias existence in European football leagues. While the model proved WNBA betting market to be inefficient, the betting simulations could not approve it since their results were statistically insignificant.

Finally, Goddard and Asimakopoulos (2004) have analyzed the English football betting league, as they implemented an ordered probit regression model with various explanatory variables, such as a distance that teams had to travel or significance of each game for championship. Their model was able to present a strategy that yielded positive or zero returns (before taxes) throughout four seasons, confirming that the market was inefficient; however, their findings agree with Gray and Gray's (1997), as they state that "inefficiencies in the bookmakers' prices have diminished over time".

2.4. Cognitive biases

The presence of various systematic biases in gambling markets is one of the main reasons why these markets have been examined for the past several decades. Since bookmakers are profit-maximizing institutions, they seek for the highest margin, therefore, creating misprices for odds instead of trying to stabilize the demand by finding a balance between them (Radiul & Janusonis, 2020; Croxson & James Reade, 2013). According to Kipkorir (2017), the cognitive biases have their patterns which punters are subject to, such as overconfidence, which goes along with risk-loving behavior, representativeness, or confirmation bias, "whereby one seeks selective information that supports their own beliefs", which might contribute to the favoritism of home team, also known as the home bias. Therefore, this section inspects existing behavioral biases in the economic literature about betting markets. There are three main types of cognitive biases that bettors are subject to, namely home-away, favorite-longshot, and most-likely/least-likely biases (Vlastakis et al., 2009; Constantinou & Fenton, 2013; Gray & Gray, 1997; Radiul & Janusonis, 2020).

As just mentioned, one of the main behavioral biases present in the betting markets is favorite-longshot bias. According to Ottoviani and Sorensen (2008), the first time favorite-longshot bias appeared in the literature was when Griffith (1949) noticed that betting on favorite horses results in higher payouts than trying luck on longshot horses. Vlastakis et al. (2009) define this particular bias as "an observed tendency for favorites to be underbet and longshots to be overbet with respect to their objective probabilities of occurrence". Here, favorites refer to a team that has more chances to win in a probabilistic sense (probability >50%), and longshots refer to the opposite. This bias exists when betting on favorites yields higher expected returns (on average) than wagering on the longshots (Sauer, 1998; Radiul & Janusonis, 2020). There can be different reasons why punters are highly subject to this bias, one of which is risk-loving behavior, which is explained by "preference of the bettor in backing risky outcomes" (Constatinou & Fenton, 2013). Assuming that the bookmakers are aware of favorite-longshot bias existence in the market, they know that punters would bet on underdogs even if they represented lower odds rather than risk their capital on favorites with more-than-fair odds in order to achieve a significantly higher payout. Thus, the most popular scenario of how the bookmakers exploit this bias is by adjusting favorite odds to become more appealing, and longshot odds to look more repellent to the wagerers (Constaninou & Fenton, 2013; Radiul & Janusonis, 2020). This instance was found by numerous researchers in various sport fields, such as horse races (Thaler & Ziemba, 1998; Deschamps & Gergaud, 2008), soccer (Paton & Vaughan Williams, 1997; Vlastakis et al., 2009), or college basketball (Berkowitz, Depken & Gandar, 2017). On the other hand, there were some rare cases of the opposite phenomena in tennis (Forrest & McHale, 2007). To sum up, according to Newall and Cortis (2021), favorite bias has more frequently appeared in sports with only two potential outcomes (win or lose), while longshot bias exists in some sports with multiple ending possibilities.

Another strong behavioral bias in the betting markets is the home-away bias. Similar to favorite-longshot, it occurs when punters tend to overestimate the chances of a team playing at home because there is a common belief that they have a home-field advantage. This phenomenon occurs not only in the gambling markets – Strong and Xu (2003) in their survey-based research have noticed the equivalent bias in typical financial markets worldwide, where investors tend to "show significant relative optimism towards their home equity market". In betting markets, a home-away bias exists if betting on home team yields higher average returns than betting on away team (Constantinou & Fenton, 2013; Radiul & Janusonis, 2020). According to Stanek (2017), there exists various factors why punters are subject to this bias. First, he points out that aspect of popularity could be a reason why the home team is prioritized since many bettors are sports fans themselves and tend to support their national teams. Moreover, familiarity plays a big role in this instance, as some researchers argue that bettors tend to overestimate familiar teams' chances to win (Stanek, 2017; Pachur & Biele, 2007). The home-away bias exists in variety of sports in different countries, for instance, European football tournaments (Vlastakis et al., 2009; Radiul & Janusonis, 2020) or Czech ice-hockey betting market (Stanek, 2017). Radiul and Janusonis (2020), who analyzed Latvian higher football league, are in agreement with Vlastakis et al. (2009), who found out that the home-field advantage is consistently overestimated by the bookmakers in various European football national leagues. On the other hand, Stanek (2017) provides strong evidence of home bias being present in the Czech ice-hockey betting market using logistic regressions. Finally, Constantinou and Fenton (2013) argues that the home-bias appears in the betting markets at a homogenous level of strength as the favorite-longshot bias.

Lastly, most-likely/least-likely bias is presented by Constatinou and Fenton (2013) and has the same concept as favorite-longshot bias, the only difference being that favorite-longshot bias considers only very high (e.g., 80%) or very low (e.g., 10%) probability events, while most-likely/least-likely takes into consideration all events, dividing them into two groups (probability to happen equal to more or less than 50%). They find that this bias is present in English and Italian national football leagues, and so conclude Radiul and Janusonis (2020) about the Latvian football league.

Thus, there are numerous existing research analyses on various fields of sports betting markets' efficiency, which are based on a similar theory and market efficiency concepts. The majority of them find at least some behavioral biases and inefficiencies by creating econometric models and trying to accurately forecast outcomes and probabilities of sports events based on historical data. However, EuroLeague, which is considered to be the second-best basketball league worldwide after the NBA, has not been examined yet; therefore, in this analysis, I represent my thoughts and findings about EuroLeague betting market efficiency.

3. METHODOLOGY AND DATA

3.1. Data description

Data needed for this research project was obtained from two websites (www.euroleague.net and www.oddsportal.com). The former is the official webpage of the Euroleague, which provides various statistics from all the seasons of its existence, while the latter is sports odds comparison service. OddsPortal.com collects and shares the archives of betting odds of more than 80 bookmakers for more than 25 different sport games. The data was collected from both websites using web scrapping[1]. Since the idea was to obtain the data from the last 10 seasons of Euroleague, the first record of data comes from October 17th, 2011, while the last one was recorded on February 22nd, 2021 – just before cancelling the season due to the Covid-19 pandemic. However, after the methodology was chosen, the latter season and playoffs of all the seasons were excluded from the dataset in order not to have any unnecessary biases. Thus, the data has 1,843 observations of games statistics and betting odds from season 2011-2012 to season 2019-2020. Moreover, as this paper concentrates on Lithuanian betting market, the only bookmaker that operates in Lithuanian gambling area and has had data available for all these seasons was analyzed. Since there are two possible endings to a basketball game – either a home win or away win – therefore, every match has two betting odds for it, which makes 3,686 odds in total. Finally, the initial data consists of the names of both teams playing, the date, final score, home, and away odds.

Table 1

Summary of match statistics relating to game scores

	Obs.	Mean	Std. Dev.	Min	Max
Winning score	1843	83.51	9.86	54	116
Losing score	1843	72.6	9.89	41	106
Winning margin	1843	10.91	8.25	1	46

This table is based on 1,843 Euroleague games from 2011 to 2020. The number of points scored by the winning and losing teams corresponds to winning and losing scores, respectively, and winning margin is the difference between them. Created by the author on the basis of the research.

[1] The software used for web scrapping was Python and Selenium.

After obtaining the initial data, some additional statistics for each season were separately collected for constructing more complex models. Firstly, I obtained a total of every team's winning margin (points scored by the opponent team subtracted from points scored by the team of record) and calculated its average (winning margin divided by games played in the current season). Additionally, I took out the records of every team's performance (total of wins and losses) and calculated their winning percentage by season (total of wins divided by games played). During any basketball match, there is always team A and team B, thus, for every game one of the teams has been defined as the *team of record*, depending on a chosen method. The first method of choosing the team of record was completely random. The second way of defining the team of record was picking the team which plays at its homecourt; and, finally, the team of record might as well be the favorite team (the one which has a smaller odd set by the bookmaker before match). The last two methods might uncover some conditional biases, such as underdogs or away teams obtaining too high predetermined odds, while the first method of choice shall not reveal any bias due to being completely random.

Table 2

Summary statistics of odds dataset

	Obs.	Mean	Std. Dev.	Min	Max
Home odds	1843	2.69	2.18	1.01	17
Away odds	1843	2.8	2.4	1.01	17.5
Favorite odds	1843	1.39	0.24	1.01	1.91
Underdog odds	1843	4.12	2.6	1.9	17.5
Random odds	1843	2.69	2.18	1.01	17
Payout for bookmaker	1843	0.95	0.004	0.91	0.96

Descriptive statistics of odds are shown for home, away, favorite, and underdog teams. Underdog team is the one with a higher predetermined odd. Payout for bookmaker is calculated as follows: $1 / (1 / (home\ odds) + 1 / (away\ odds))$. If it is <1, the bookmaker is making profit. Created by the author on the basis of the research.

3.2. Mathematical framework

Since the goal is to check whether Lithuanian betting market is efficient, one needs to have a better understanding of how the inefficiencies occur. In order to achieve it, firstly, we need to describe how the betting markets function. This paper concentrates on fixed odds, therefore, their creation process is described in this part.

Odds do not appear only in gambling – they are used in statistics as well. However, the two quite differ because statisticians do not need a profit margin. Generally, statistical odds show a true probability of an event; on the other hand, gambling odds do not represent that, and they can be described as the ratio of profit won to the stake (Radiul & Janusonis, 2020). There are a few types of gambling odds, such as decimal, American, or already described fractional odds. This research paper analyzes decimal odds, which represent a profit for a bettor for every 1 euro placed. For instance, if the bettor wagers 3 euros on 2.32 decimal odd, the potential payout (in case bettor wins) can be computed by multiplying 3 by 2.32 and deducting the initial value placed (3 euros) from it, making it 3.96 euros the realized profit.

Moreover, bookmakers set the betting odds themselves. As mentioned before, bookmakers, in order to maximize their profit, do not form odds that represent the true probability of an event, and in this way, they receive additional money. This additional money they want to secure is known as a commission in the betting world, also called as an over-round or a profit margin, which can be defined as “the margin by which the sum of the probability odds of the total outcomes exceeds one” (Constantinou and Fenton, 2013; Radiul & Janusonis, 2020). This means that the bookmakers form the odds fairer to themselves rather than the clients. The only way that the odd can be defined as fair is when the sum of likelihoods of an event (the sum of winning probabilities for both teams) is equal to 1. Furthermore, gambling odds usage differs in theory and in reality. In theory, as said earlier, the odds shall serve as true probabilities and, thus, a bookmaker shall receive a payoff exactly as the over-round (Franck et al., 2013; Radiul & Janusonis, 2020). On the other hand, in reality, bookmaker’s profit margin is “just an approximation of a bookmaker’s expected profit” since the bookmakers’ payout is not the same for every bettor and does not depend on the outcome of the game (Constantinou and Fenton, 2013). As Levitt (2004) and Vlastakis et al. (2009) have discussed, bookmaker’s expected gain on an event with *n* outcomes might be calculated as:

$$E(M) = 1 - \sum_{i=1}^{n} P_i \times w_i \times d_i \qquad (1)$$

In the equation above, the expected payoff is denoted by M, and the independent variables are probability of every outcome - P_i, the percentage of bets on each outcome – w_i, and quoted odds - d_i. It is not known how exactly the bookmakers set the odds but there are a few ways to do it. The first method would be by predicting the actual outcome of the match, which might depend on various factors, such as a home field, winning and losing streaks, and others. Secondly, the bookmakers might do it by predicting what the distribution of bets on each outcome would be (Levitt, 2004). Nevertheless, both of just mentioned methods could be used.

Furthermore, not all the data needed to calculate the expected margin for bookmakers is public. Although anyone can see the actual odds, the percentage of bets on each outcome is not publicly available. Since we do not have the data needed for equation (1), the expected payoff – M – cannot be computed. Even so, Vlastakis et al. (2009) assume equal distribution of the bets and that odds are based on true probabilities. However, it would mean that the potential bookmaker margin would be equal to zero. This is the reason why the fair odds are larger than the actual ones (to gain some payoff for bookmakers), and, therefore, the actual odds "do not correspond to true probabilities but to somewhat larger implied probabilities" – P_i^*. Thus, according to Vlastakis et al. (2009), the expected implied gain – M^* might be calculated as:

$$E(M^*) = \left(\sum_{i=1}^{n} P_i^*\right) - 1 = \left(\sum_{i=1}^{n} \frac{1}{d_i^*}\right) - 1 \qquad (2)$$

Finally, Gray and Gray (1997) states that the betting markets are inefficient if odds set by the bookmakers do not correspond to the sum of actual likelihoods of outcomes of an event. In this case, the odds represent systematic bias.

3.3. Models

3.3.1. Benchmark model

We cannot examine whether betting markets efficiently incorporate data that is not publicly available; however, it is possible to analyze the efficiency having the public data only. Using Probit and Logit models and having obtained the dataset described previously, I have examined if any public data is useful in forecasting sports odds, and, therefore, if Lithuanian betting market is operating efficiently. In this research, probit and logit models were chosen for a more comprehensive analysis.

Since the same dependent and independent variables were used for each of the models, upcoming equations and model descriptions are identical for both. The main difference between them is that the probit model is based on cumulative distribution function of the standard normal distribution, while the logit model uses a logistic function. At first, a linear model estimated using OLS was also considered in the analysis but it was excluded as such econometric approach is not suitable when the dependent variable is binary. Moreover, it works better with forecasted point spread as a regressor, which cannot be calculated having the implied probabilities to win only. Even if this information were available and used with OLS, one could not draw robust conclusions because events with a larger winning margin have a tendency for a greater impact on the slope coefficient (Gray and Gray, 1997). Since it is irrelevant by how many points a team of record beat the forecasted spread in betting markets, therefore, rather a discrete choice variable, defined below, was used as a dependent variable:

$$Y_i = \begin{cases} 1 & if\ predicted\ favorite\ won \\ 0 & otherwise \end{cases}$$

According to Gray and Gray (1997), in order for betting markets to be efficient, there should not be any observable information that might assist to predict Y_i in a statistically significant sense. Before introducing new variables, below is the benchmark model for both probit and logit:

$$Y_i^* = \alpha_0 + \varepsilon_i \qquad (3)$$

where α_0 is the constant and Y_i^* is the dependent variable which is defined above. This simple model with dependent variable and an intercept only is the baseline of all the following models. The likelihood value of this model is then used to construct likelihood ratio test (LRT) which is presented at the end of this section.

3.3.2. Augmented models

Furthermore, home and favorite dummies are included in the following models to test for home- and favorite-longshot-biases:

$$Y_i^* = \alpha_0 + \alpha_1 Home_i + \alpha_2 Favorite_i + \varepsilon_i \qquad (3.1)$$

where the team of record is defined randomly. In both probit and logit models, the dependent variable Y_i is equal to 1 if Y_i^* has a positive value, and equal to 0 otherwise, where Y_i is defined above and

$$Home_i = \begin{cases} 1 & if\ the\ team\ of\ record\ is\ playing\ at\ home\ field \\ 0 & otherwise \end{cases}$$

$$Favorite_i = \begin{cases} 1 & if\ the\ team\ of\ record\ is\ predicted\ favorite \\ 0 & otherwise \end{cases}$$

These augmented probit and logit models examine whether there is any statistical significance between the dependent variable and two dummies - home and favorite; in other words, I analyze the relationship and significance between the (1) outcome of the match, (2) home-field advantage and (3) favorite status. Home bias means that bettors tend to overvalue home team's potential to win and there are a few ways how the bookmakers try to exploit it, which are described in section **4.3.2**.

Furthermore, I include one more independent variable to the regression, which is an interaction of home and favorite dummy variables, calculated by multiplying them. It shows whether there is a possible home-favorite bias. Additionally, it is also included in the betting simulation, which not only represents average profits on betting on favorites that are playing at home, but also checks if there is any existence of home-underdog, away-favorite or away-underdog biases. Thus, now the third version of the augmented probit-logit models look like this:

$$Y_i^* = \alpha_0 + \alpha_1 Home_i + \alpha_2 Favorite_i + \alpha_3 Home_i \times Favorite_i + \varepsilon_i \quad (3.2)$$

where the variable $Home_i$*$Favorite_i$ is equal to 1 if the team of record is playing at home **and** is a predicted favorite, and it is equal to 0 otherwise. Augmenting the models even further, I have included six more explanatory variables. As described in the data section, these variables are a total of every team's winning margin (total of points scored by their opponents deducted from total of points scored by the team of record), its average (winning margin divided by games played in the current season) and every team's overall performance, computed by taking a number of wins (of one team of every season separately) and divided by games played. According to Gray and Gray (1997), if at least one variable, which can be obtained from public data, is statistically significant in these regressions, a conclusion that betting markets are inefficient is drawn because it reveals that markets might not use all the information available for the process of forming odds. Therefore, having split my created variables, which are computed from data obtained from the official public webpage of Euroleague, into two groups, I checked if they could help to detect whether the Lithuanian betting market is efficient:

$$Y_i^* = \alpha_0 + \alpha_1 Home_i + \alpha_2 Favorite_i + \alpha_3 Home_i \times Favorite_i + \alpha_4 wp1_i + \alpha_5 wp2_i + \alpha_6 apd1_i + \alpha_7 apd2_i + \varepsilon_i \quad (3.3)$$

and

$$Y_i^* = \alpha_0 + \alpha_1 Home_i + \alpha_2 Favorite_i + \alpha_3 Home_i \times Favorite_i + \alpha_4 wp1_i + \alpha_5 wp2_i + \alpha_8 pd1_i + \alpha_9 pd2_i + \varepsilon_i \qquad (3.4)$$

where *pd1* and *pd2* represent the total point difference of one season, *apd1* and *apd2* – an average point spread, and *wp1* and *wp2* – winning percentages of one season for the first and the second teams, respectively. As mentioned before, if any of these variables showed significance in the regression, the market is not operating efficiently.

Additionally, I included a dummy for every season separately to all of the models above (excluding the benchmark model), to check whether the market is efficient given that the seasonal effects are accounted for. A seasonal dummy takes a value of 1 if it is the season of record, and 0 otherwise. The α coefficients for them begin with α_{10} and end with α_{17} for seasons from 1 to 8, respectively (*season 1* representing the season of 2011-2012, while *season 8* – the season of 2018-2019). Finally, a so-called likelihood ratio test was implemented to test that all the coefficients except the intercept are jointly significant. If the null hypothesis is rejected, this would suggest that there is at least one coefficient of the model that is statistically significant which in turn suggests statistical evidence against market efficiency. The results of all the models are represented and the conclusions are drawn in section 5.

3.4. Odds biases

Since bookmakers are profit-maximizing organizations, they try to maximize their profit in any way possible. One of the cases of capturing more gain is already described above – odds set by the bookmakers do not represent the true probabilities of an event, which allows them to have a larger payoff. Furthermore, another example of how bookmakers try to maximize their profit is by taking behavioral biases into account. There are a few main types of bettors' behavioral biases, such as home-away, favorite-underdog, or least-likely/most-likely, which are defined in the sections below.

3.4.1. Home bias

First of all, home bias exists not only in betting markets. It typically appears in financial markets when there is a tendency for investors to focus their equity and get into the domestic market. In betting markets, this phenomenon is very similar and defined as "bettor's tendency to bet predominantly on their home team" (Stanek, 2017). Since bettors have the incentive to wager money and the true value of bets become known very quickly (most of the bets are made on matches that take place in the near future), the gambling market is a proper environment for analyzing some typical behavioral biases, such as the home bias itself.

Home bias, also known as a home-away bias, occurs when punters tend to gamble more on the home team. According to Stanek (2017), the bookmakers might exploit this bias in a few ways. First of all, only the odds for the home team might be decreased. This way, bettors would be more willing to bet on the away team, and, as represented in the section of results, leading to lower payoffs for the punters. The second way would be to balance the stakes – not only to decrease odds in favor of the home team but also to increase the odds in favor of the away team. Finally, in both cases, there is a lower payout for the punters if they are subject to the home-away bias.

This paper examines the home-away bias in two different ways: one of them is represented above (equations 3.1 to 3.4), where a home dummy is included in both probit and logit regressions, taking a value of 1 if a team of record is playing at home, and 0 otherwise. The other way to check for the home bias is the same as in the section below – I simulated a bet of 1 euro on both home and away wins. According to Constantinou and Fenton (2013), a home-away bias exists when the profit of betting on the home team exceeds the cumulative payouts for betting on the away team. The results and conclusions are represented in the results section, and the betting simulation is described more in-depth in the paragraph below.

3.4.2. Most-likely/least-likely and favorite-longshot biases

The most-likely/least-likely and favorite-longshot biases represent the cases when bettors tend to overestimate the underdogs and underestimate the favorites. Since bets on low odds (the events that have a high probability) generate higher profitability than bets on high odds (the events that have a low probability), and there is a tendency of risk-loving behavior (especially in the betting market), the favorite-longshot bias exists (Constantinou and Fenton, 2013). For instance, the best team of the tournament is playing at home against the worst team. The home team has a true probability of 90% to win the match. Even if it seems like a 'safe' bet (suppose the bookmakers set a fair odd of 1.1 for the home team), the bettors are more expected to place their bets on the away team, which in this instance could have a less than fair odd on it, say 13. It means that if a punter bets 10 euros on the home team, he is expected to win only 1.1 euros, and if the bet of 1 euro is placed on the underdog, an expected profit would be 13 euros. In the literature, it is widely believed the bookmakers are well-aware of this behavioral bias and exploit it “by offering more-than-fair odds for ‘safe’ outcomes, and less-than-fair odds for ‘risky’ outcomes (Constantinou and Fenton, 2013).

The two systematic biases of this section are very similar. Their analysis starts with the actual odds being converted to implied probabilities. For the favorite-longshot bias, I chose the upper and the lower boundaries of event probabilities, considering matches with high or low likelihoods only. Similar to what Constantinou and Fenton (2013) did in their examination on behavioral biases for the field of football, I have split high probability events into 3 groups – more than 60%, 70%, and 80% chance of happening, and low probability events – less than 10%, 15%, and 20% - into 3 groups as well. Additionally, following Radiul and Janusonis (2020) methodology, I have included one more group of high probability of 90% to check whether it is possible to earn a profit by betting on clear favorites only. These 7 groups were examined in a betting simulation where, depending on the type of group, the same amount of money was placed on favorites or longshots. The favorite-longshot bias occurs when the average cumulative returns from betting on underdogs are notably lower than betting on favorites (Radiul and Janusonis, 2020). Similarly, if the average profit of betting on most-likely events outweighs the generated returns of betting on the least-likely events, it is a sign of most-likely/least-likely bias presence in the market. This behavioral bias is somewhat similar to the favorite-longshot, only here there are two groups, one of which is a favorite of the match, having more than 50% chance of winning, and another is the underdog (less than 50% chance of winning). The same betting simulation was conducted for both instances to check whether there are behavioral biases in the Lithuanian betting market. The results and graphs are presented in the section below.

4. RESULTS AND RESEARCH FINDINGS

4.1. Data overview

The summary tables in the data section, Tables 1 and 2, might suggest a few useful insights before going further into the analysis. First, the payout for bookmaker indicates that the average profit of my chosen bookmaker was 5 per cent during seasons 2011/12 – 2018/19. Furthermore, the winning margin (Table 1) was close to 11, suggesting a possible favorite-longshot bias in the market. The averages of favorites (1.39) and underdogs (4.12) odds in Table 2 are consistent with the potentiality of the bias. Talking about the home-away bias, the home and away scores were not included. However, I calculated that the home team has scored an average of 79.9 points per game, while the away team - 76.2 points per game. In basketball, a difference of almost 4 points might not look much; on the other hand, it could be the first sign of a possible home-away bias. Every now and then the bookmakers might insure against this potential bias, as it is widely discussed in the literature that they have a strategy of forming lower than fair odds for the home team and higher odds for the away team (Levitt, 2004). Table 2 is evidence of that, showing that even if it seems not significant, but there is a difference between home (2.8) and away (2.69) odds on average. It could raise a theory that the bookmakers, which are profit-maximizing institutions, are aware of presence of the biases, and, therefore, "hedge against the possible risks to minimize the losses, and, on the contrary, skew the odds to realize more gains" (Radiul and Janusonis, 2020).

4.2. Statistical models of Market Efficiency

To have a more comprehensive analysis of the efficiency of the EuroLeague betting market, several statistical tests were implemented to check whether different publicly available data could assist to forecast outcomes of the events. According to Gray and Gray (1997), a gambling market is inefficient if there is at least one variable computed from public data that is possibly useful for predicting how a match could end. As described in the data section, the independent variables for both probit and logit regressions are (0) constant, (1) home dummy, (2) favorite dummy, (3) an interaction of favorite and home dummies, (4 and 5) seasonal winning percentages for both teams, (6 and 7) average seasonal point differences, and (8 and 9) actual seasonal point differences.

Table 3

Results of the probit models

Parameter	1st model	2nd model	3rd model	4th model	5th model
Intercept α_0	0.54	0.51***	0.61***	-0.05	0.69**
	(-0.03)	(0.05)	(0.05)	(0.34)	(0.27)
Home dummy α_1		0.03	-0.26**	-0.32**	-0.32**
		(0.07)	(0.09)	(0.10)	(0.10)
Favorite dummy α_2		0.03	-0.28**	-0.33**	-0.29**
		(0.07)	(0.09)	(0.11)	(0.11)
Favorite and Home dum. α_3			0.60***	0.62***	0.60***
			(0.13)	(0.13)	(0.13)
Winning percentage 1 α_4				0.53	0.06
				(0.45)	(0.37)
Winning percentage 2 α_5				0.84	-0.15
				(0.45)	(0.37)
Average point diff. 1 α_6				-0.02	
				(0.02)	
Average point diff. 2 α_7				-0.03	
				(0.02)	
Point difference 1 α_8					0.00
					(0.00)
Point difference 2 α_9					0.00
					(0.00)
LRT		0.71	8e-05***	0.00***	0.0012**

The first model corresponds to the benchmark model, the second includes home and favorite dummies, the third is the same as the second only with additional variable of an interaction of favorite and home dummies, while the fourth and the fifth models also have seasonal winning percentages and either an average or an actual seasonal point difference. Standard errors are in parentheses. Created by the author on the basis of the research.

The first model is the benchmark model, having an intercept as the only regressor. The second model is augmented with a few more variables, namely home and favorite dummies. The likelihood ratio test (LRT) of this model yields a statistic of 0.71, which is insignificant at any usual level. It is the only model where home and favorite dummies show no significance on the dependent variable. However, the third model is introduced, which is the same as the latter but with one additional dummy – an interaction of favorite and home dummies. This model along with the upcoming fourth and fifth models are jointly significant. The results of the third model represent the first two dummies to be significant at 0.05, while their interaction – at 0.01 level. The negative signs for α_1 and α_2 suggest that away-favorites are more likely to achieve a win (as predicted) than the home-favorites. However, it opposes the results of α_3 which show the opposite at a higher statistical significance level, leading to a conclusion that the home-favorites are most-likely to secure the win, as the betting simulation approves of this hypothesis.

Table 5

Results of the seasonal probit models

Parameter	6th model	7th model	8th model	9th model
Intercept α_0	0.34***	0.43***	-0.15	0.44
	(0.09)	(0.09)	(0.35)	(0.19)
Home dummy α_1	0.04	0.27**	-0.31**	-0.30**
	(0.07)	(0.09)	(0.10)	(0.10)
Favorite dummy α_2	0.03	-0.30**	-0.24*	-0.23*
	(0.06)	(0.09)	(0.11)	(0.11)
Favorite and home dummies α_3		0.62***	0.64***	0.62***
		(0.13)	(0.13)	(0.13)
Winning percentage 1 α_4			0.14	-0.35
			(0.45)	(0.37)
Winning percentage 2 α_5			1.01*	0.31
			(0.46)	(0.37)
Average point difference 1 α_6			-0.01	
			(0.02)	
Average point difference 2 α_7			-0.03*	
			(0.02)	
Point difference 1 α_8				0.00
				(0.00)
Point difference 2 α_9				0.00
				(0.00)
Season 1 α_{10}	0.28*	0.30*	0.28*	0.31*
	(0.13)	(0.13)	(0.14)	(0.13)
Season 2 α_{11}	0.13	0.14	0.12	0.14
	(0.12)	(0.12)	(0.12)	(0.12)
Season 3 α_{12}	0.32*	0.33**	0.31*	0.33**
	(0.12)	(0.12)	(0.12)	(0.12)
Season 4 α_{13}	0.17	0.19	0.17	0.19
	(0.12)	(0.12)	(0.12)	(0.12)
Season 5 α_{14}	0.12	0.14	0.13	0.13
	(0.13)	(0.13)	(0.13)	(0.13)
Season 6 α_{15}	0.09	0.07	0.06	0.07
	(0.12)	(0.12)	(0.12)	(0.12)
Season 7 α_{16}	0.21	0.2	0.19	0.2
	(0.15)	(0.16)	(0.16)	(0.15)
Season 8 α_{17}	0.24*	0.24*	0.23*	0.24*
	(0.12)	(0.12)	(0.12)	(0.12)
LRT	0.39	0.0006***	0.0008***	0.002**

The models above correspond to the 2nd-5th models from Table 4, with seasonal dummies included. LRT statistic is calculated using log-likelihood values of two models. Statistical significance and standard errors are denoted in the same way. Created by the author.

Furthermore, results from the fourth and the fifth models (where variables are jointly significant since the log-likelihood values for both are close to 0) agree with the findings from the previous models, representing almost identical numbers for the three dummy variables. Nevertheless, these models have additional regressors, namely winning percentage, point difference and average point difference, all of which are seasonal; however, none of them showed significance at any usual level.

On the other hand, seasonal models presented in Table 5 suggest that the winning percentage and average point difference (for the second team) could be useful information in predicting the outcome of the match, both showing a significance at 10% level[2]. Again, any existing data which might help to forecast the outcome of an event present evidence of market inefficiency (Gray and Gray, 1997). These results, however, might be ambiguous since these variables are calculated only at the end of each season; nevertheless, one might always compute an actual winning percentage or average point difference, which, according to the results, would be beneficial for the bettors.

Moreover, as expected, the results for home and favorite dummies are almost identical as for the previous models, with two exceptions only. First, the coefficient on the home variable is positive, concurring with the idea that home-favorites are more-likely to achieve a victory than the away-favorites. Second, the statistical significance levels for α_3 show significance at 10% level for models 8 and 9, unlike the corresponding results from model 7 and the models presented in Table 4. Nevertheless, the negative signs and sufficient significance level agree with the previous results.

Lastly, the seasonal dummies represent that the degree of EuroLeague betting market inefficiency varies across different seasons. Coefficients for seasons 1 (2011/12), 3 (2013/14), and 8 (2018/2019) were positive and showed significance at either 5% or at 10% significance levels for every seasonal model, leading to a conclusion that the market was even more inefficient during them. Finally, the likelihood ratio test[3] proved that variables for models 7-9 were jointly significant, while model 6, which yielded the opposite results from the others, was insignificant.

Going further, the same models were run through logistic models' family to check whether conclusions drawn from the results above are consistent. These results are presented in Tables 6 and 7.

---

[2] Statistical significance is denoted as following: * corresponds to p<0.1, ** - p<0.05, *** - p<0.01.

[3] LRT statistic is calculated using log-likelihood values of two models.

Table 6

Results of the logit models

Parameter	**1st model**	**2nd model**	**3rd model**	**4th model**	**5th model**
Intercept α_0	0.88***	0.83***	0.99***	-0.20	1.03*
	(0.05)	(0.08)	(0.09)	(0.57)	(0.44)
Home dummy α_1		0.06	-0.43**	-0.49**	-0.48**
		(0.11)	(0.15)	(0.16)	(0.15)
Favorite dummy α_2		0.04	-0.46**	-0.37*	-0.36*
		(0.11)	(0.15)	(0.18)	(0.18)
Favorite and Home dum. α_3			0.99***	1.03***	1.01***
			(0.21)	(0.22)	(0.22)
Winning percentage 1 α_4				0.41	-0.62
				(0.74)	(0.61)
Winning percentage 2 α_5				1.85*	0.46
				(0.76)	(0.62)
Average point diff. 1 α_6				-0.02	
				(0.02)	
Average point diff. 2 α_7				-0.05*	
				(0.03)	
Point difference 1 α_8					0.00
					(0.00)
Point difference 2 α_9					0.00
					(0.00)
LRT		0.71	8e-05***	0.00***	0.0011**

The same variables were used as for the probit family for every model, respectively. LRT statistic is calculated using log-likelihood values of two models. Statistical significance and standard errors are denoted in the same way. Created by the author.

The results for all of the logit models were rather consistent with the corresponding results from the probit models, again showing strong evidence for weak-form market inefficiency. Both home and favorite dummies are negative and statistically significant, while their interaction is positive at even higher significance level of 1%, as this agrees with the previous conclusion of home-favorites having the edge over the away-favorites. Moreover, winning percentage and average point difference were positive and negative, respectively, and significant at 10% level, suggesting additional evidence supporting the hypothesis of EuroLeague gambling market being inefficient. Also, the same three seasons (2011/12, 2013/14 and 2018/19) showed significance, which concurs to the idea that the degree of market inefficiency differs across different seasons. Lastly, the likelihood ratio test suggested that the variables of models 2 and 6 were insignificant, while the others remained jointly significant, which, once again, proved that the market is experiencing a weak-form inefficiency.

Table 7

Results of the seasonal logit models

Parameter	6th model	7th model	8th model	9th model
Intercept α_0	0.55***	0.71***	-0.30	0.69
	(0.15)	(0.15)	(0.58)	(0.48)
Home dummy α_1	0.06	-0.44**	-0.51**	0.49**
	(0.11)	(0.15)	(0.16)	(0.16)
Favorite dummy α_2	0.04	-0.47**	-0.37*	-0.35*
	(0.11)	(0.15)	(0.18)	(0.18)
Favorite and home dummies α_3		1.03***	1.06***	1.03***
		(0.22)	(0.22)	(0.22)
Winning percentage 1 α_4			0.22	0.57
			(0.74)	(0.62)
Winning percentage 2 α_5			1.7*	0.53
			(0.76)	(0.62)
Average point difference 1 α_6			-0.02	
			(0.03)	
Average point difference 2 α_7			-0.05*	
			(0.03)	
Point difference 1 α_8				0.00
				(0.00)
Point difference 2 α_9				0.00
				(0.00)
Season 1 α_{10}	0.47*	0.51*	0.47*	0.51*
	(0.22)	(0.22)	(0.22)	(0.22)
Season 2 α_{11}	0.22	0.23	0.21	0.22
	(0.20)	(0.20)	(0.20)	(0.20)
Season 3 α_{12}	0.52*	0.54**	0.52*	0.55**
	(0.20)	(0.21)	(0.21)	(0.21)
Season 4 α_{13}	0.28	0.32	0.29	0.32
	(0.19)	(0.20)	(0.20)	(0.20)
Season 5 α_{14}	0.20	0.23	0.21	0.22
	(0.21)	(0.21)	(0.21)	(0.21)
Season 6 α_{15}	0.15	0.12	0.1	0.11
	(0.2)	(0.2)	(0.2)	(0.2)
Season 7 α_{16}	0.36	0.33	0.32	0.33
	(0.26)	(0.26)	(0.26)	(0.26)
Season 8 α_{17}	0.40*	0.4*	0.38*	0.4*
	(0.20)	(0.2)	(0.2)	(0.2)
LRT	0.39	0.0006***	0.0007***	0.002**

The same variables were used as for the probit family for every model, respectively. LRT statistic is calculated using log-likelihood values of two models. Statistical significance and standard errors are denoted in the same way. Created by the author.

4.3. A betting simulation

To further analyze the appearance of behavioral biases in the betting markets, a betting simulation was performed to evaluate the differences in average cumulative returns by placing a 1 euro bet on all of the groups that are present in Table 7. First, as Vlastakis et al. (2009) stated, the favorite-longshot bias exists if the average returns from betting on the most-likely outcomes (events with probability larger than 50 per cent) significantly exceed the returns gained from betting on the least-likely outcomes (probability smaller than 50 per cent). The results shown in Table 7 are strong evidence of this bias being present in the Lithuanian betting market, as consistently betting on favorites yields a significantly higher profit (an average -0.05 euro) than betting on longshots (an average of -0.11 euro). Furthermore, I examined various probability groups of favorites and longshots; the results clearly indicate that the loss of betting on longshots increases as the probability of an event decreases, and it is getting closer to zero when the probability of favorites increases.

Table 8

A betting simulation

Condition	Number of bets	Profit per bet
Bet on longshot (Pr<0.1)	46	-0.4335
Bet on longshot (Pr<0.15)	131	-0.2769
Bet on longshot (Pr<0.2)	218	-0.2688
Bet on favorite (Pr>0.6)	738	-0.0301
Bet on favorite (Pr>0.7)	464	-0.0219
Bet on favorite (Pr>0.8)	234	-0.0196
Bet on favorite (Pr>0.9)	59	-0.0231
Bet on least-likely (Pr<0.5)	906	-0.1097
Bet on most-likely (Pr>0.5)	924	-0.0478
Bet on home team	937	-0.0388
Bet on away team	906	-0.1177
Bet on home team least-likely	544	-0.0650
Bet on home team most-likely	652	-0.0265
Bet on away team least-likely	634	-0.1480
Bet on away team most-likely	272	-0.0976

This table represents the potential profits bettors would make if betting an equal amount on various outcomes, namely most and least-likely, home and away, home (away) favorites (longshots) advantage, and different probabilities of favorite-longshot biases. Created by the author on the basis of the research.

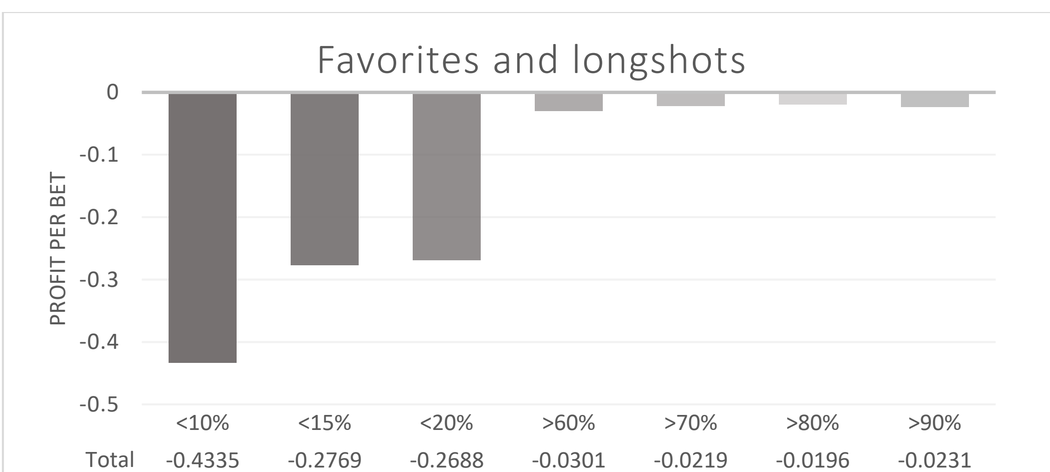
Similar to Constantinou and Fenton (2013), a bettor never breaks even with the strategy of betting on the high likelihood events only, which “implies that the bias was not strong enough to overcome the averaged bookmaking profit margin”. However, my findings are not in full agreement with Radiul and Janusonis (2020), where they found that the favorite group of 90% chance of winning yielded an almost break-even result in the Latvian football league; nevertheless, since I examined a sole bookmaker, the results might be insufficient to draw any robust conclusions about it due to a possible lack of observations. Moreover, the same simulation was executed for the outcomes of home and away wins, which once again proved the hypothesis of home-away bias being present in the market since the average return from betting on the home team significantly exceeds the return from away team winnings. However, according to Vlastakis et al. (2009), it must be considered that the odds formed by the bookmakers for both home and away teams contain a favorite-longshot bias, and that the away team is more often the underdog, as the number of observations in Table 7 is clear evidence of that. Therefore, the effects on home (away) field must be analyzed for favorites and longshots separately.

4.3.1. Favorites and longshots

Figure 1 shows the results of the 1-euro-betting simulation on different likelihood event groups for both favorites and longshots. The first three groups correspond to the low probability longshots, while the last four show the average payout for a punter if he bets on various favorite groups.

Figure 1

Results from betting simulation on favorites and longshots



This figure represents the profits of a betting simulation, where the same value of money is wagered for various probability groups of favorites and longshots. Probabilities for the longshots are less than 10%, 15%, and 20%, while more than 60%, 70%, 80%, 90% correspond to the favorites. Created by the author on the basis of the research.

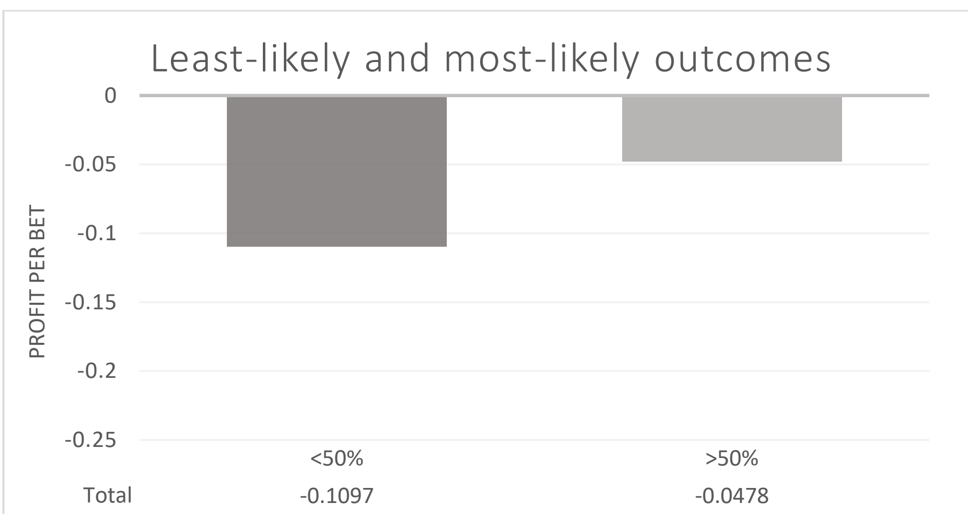
The average generated returns increase as the probability of an event increases, with the only exception of the >90 per cent group, which, nevertheless, yields a very similar result to the >80 per cent group. Hence, this behavior of average returns moving upwards indicates that top Europe's basketball league betting market is subject to the favorite-longshot bias, as the bookmakers set the attractive odds for favorites and rather repulsive odds for longshots.

4.3.2. Least-likely and most-likely outcomes

Furthermore, the results from least- and most-likely bias agree with the favorite-longshot bias. The least-likely outcomes have a less than 50 per cent implied probability, while the most-likely outcomes are the opposite. If a bettor picks the most-likely strategy, it is going to bring a bigger payout in the long-run, which is depicted in the Figure 2. A strategy to wager on the underdogs only results in a negative profit of almost 11 cents, while betting on the favorites is comparably more than two times successful; however, it does not turn out to be profitable, as a 1-euro-bet on the favorites yields a bit more than -0.05 euro payout. Nevertheless, since the most-likely outcomes yield a larger profit per bet than the least-likely outcomes, it is a proof of least-likely/most-likely bias existence in the betting market.

Figure 2

Results from betting simulation on least- and most-likely outcomes



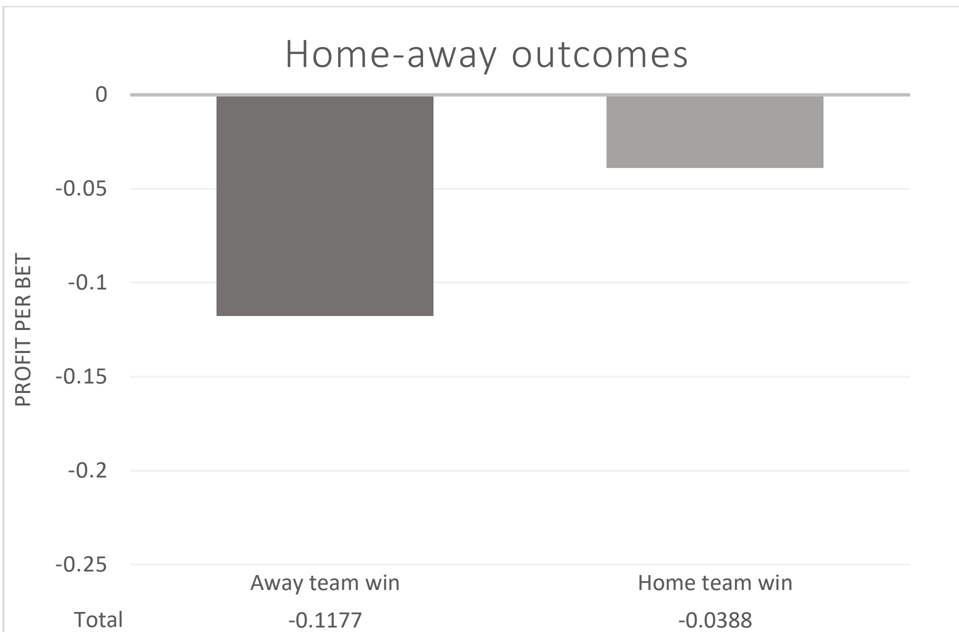
This figure shows the profits of the betting simulation for the least- and most-likely outcomes. The least-likely group of outcomes corresponds to events that have less than 50% of happening, while the most-likely represent the opposite – events with more than 50% chance of happening. Created by the author on the basis of the research.

4.3.3. Home-away outcomes

Analyzing further the behavioral biases in the betting markets, one must examine the presence of home-away bias in the market. According to Stanek (2017), an inclination to bet on the home team could arise due to various reasons, one of which is that the majority of bettors are also sports fans, which have their favorite, which, in many cases, is their national team. As said earlier, when punters are subject to the home bias, they are willing to bet on the home team even in harsher conditions, such as when they are given less favorable odds. The results of betting 1 euro on home and away teams separately are presented in the Figure 3.

Figure 3

Results from betting simulation on home or away team win outcomes

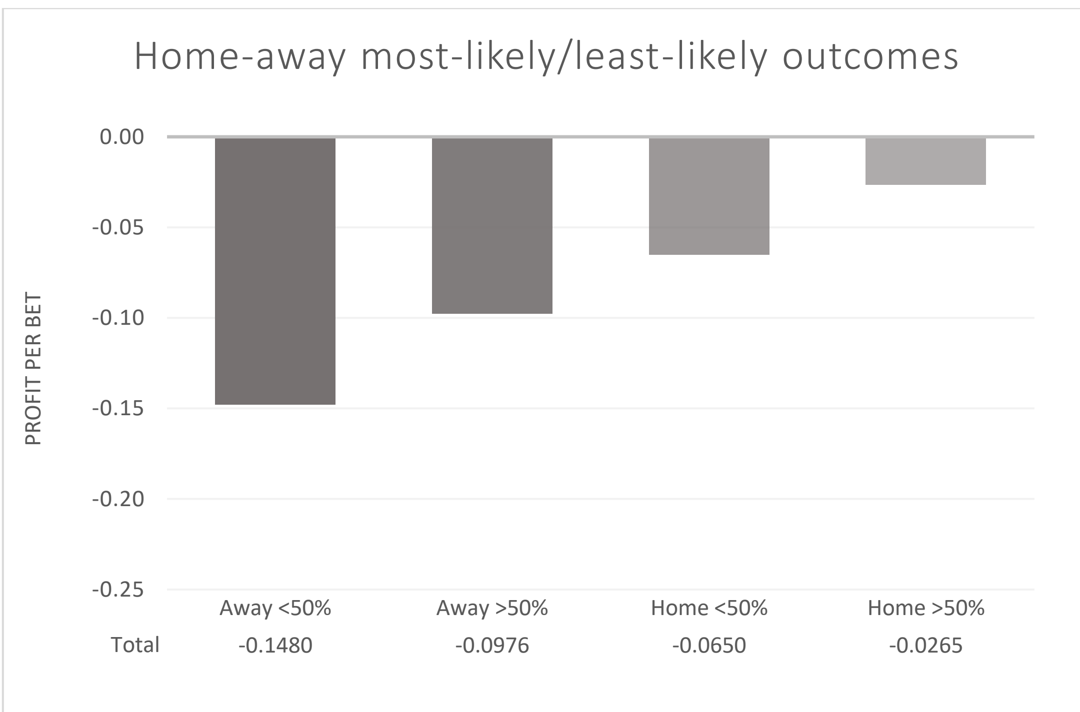


This figure depicts average profits for betting on away and home teams separately. Created by the author on the basis of the research.

As depicted above, the average generated profits for home and away teams are almost identical to the least- and most-likely situation. Since we already know from Table 7 that away teams are more likely to be underdogs, the correlation between these two biases is not a surprise. Betting on the away team only yields a negative profit of almost 12 cents, while the same strategy but for the home team only gives a punter less than -4 cents payout. In this case, the results suggest a present home-away bias in the market since the negative profit for betting on the home team exceeds the profit from the same betting value for the away team. On the other hand, examining only the profits of home and away teams might end up with ambiguous results because the home-away bias can be affected by the favorite-longshot bias. As a consequence, following Vlastakis et al. (2009) methodology, I analyzed the home field effect for most-likely and least-likely outcomes separately. Therefore, Figure 4 provides results of the analysis of home most-likely (a team of record playing at home and being a favorite to win the match), home least-likely (home team being the underdog), away most-likely (away team being the favorite), and away least-likely (away team being the underdog). Unsurprisingly, two groups with the largest number of observations were home-most-likely and away-least-likely outcomes, while the away-most-likely had more than twice observations less than the others.

Figure 4

Results from betting simulation on different outcome of the game with different probabilities



This figure represents the results of home field advantage for both favorites and longshots. The first two columns show average profits of betting on away longshots and away favorites, respectively, while the last two correspond to home longshots and favorites. Created by the author on the basis of the research.

As depicted in the figure 4, betting on both home team favorites and longshots yield better results for the punters, in comparison with betting on the away team. This is strong evidence of the home court having a consistent advantage, which, according to Vlastakis et al. (2009), indicates that the bookmakers overestimate it purposefully. The straight upward going profit line shows that punters lose the most money (an average of almost -15 cents) betting on the away longshots, while they might to nearly reach a break-even betting on the home-favorites, at least in the short-run. The results do not fully correspond to Vlastakis et al. (2009) research, where they found out that the highest profit one can make is when betting on away favorites and named this situation after the “away-favorite” bias. However, their research’s field of sport was football. Nevertheless, this analysis is in agreement with the statistically significant coefficients from the models examined in the first part of this section – there exists a strong home-favorite bias in the EuroLeague basketball betting market.

5. CONCLUSIONS

This study examines the second-biggest basketball league worldwide, EuroLeague, betting market and its efficiency. The main methodology used is econometric modeling and a betting simulation, both of which investigated if there are any present behavioral biases in the market, which help to predict the outcomes of basketball matches and search for profitable betting strategies, and, therefore, conclude about the market's efficiency.

Taking into consideration a rapid improvement of technologies, an increasing number of bookmakers and academic literature on this topic, in general, betting markets should be operating efficiently. However, this paper controverts this assumption. First, theoretically, the odds set by a bookmaker shall represent the true likelihood of different outcomes, which, in reality, was found not to be true in EuroLeague betting market.

Moreover, by setting up and examining different probit and logit regressions, various variables, such as home and favorite dummies, seasonal winning percentage and average point difference showed significance in predicting outcomes of game matches. Thus, the main finding of home-favorite bias existence in the market was presented. Since there should not be any historical data helpful to predict the outcomes, EuroLeague betting market was concluded to experience a weak-form inefficiency.

Additionally, a betting simulation was conducted to prove this hypothesis. Various groups of longshots (less than 10%, 15% and 20% probability to achieve a win) and favorites (more than 60%, 70%, 80%, 90%) presented a positive slope in profits; however, as expected, none of the strategies could yield positive results. The results of other strategies showed the most common home-away, favorite-longshot and most-likely/least-likely biases existence in the market; also, this simulation confirmed the findings of the econometric models, as punters being subject to the home-favorite bias was repeatedly found to be true. Finally, the majority of findings agree that EuroLeague betting market is weakly inefficient, as there should not be any significant historical data that may be useful in accurately forecasting outcomes of sport matches in order for the market to operate efficiently.

To sum up, my analysis contributes to the existing academic literature on betting markets' efficiency with high importance and its methodology and findings could be useful for future researchers. Its main significance is that this study was based on the top Europe's basketball league, which betting market has not been investigated before. As the idea was to concentrate on the Baltics, only one bookmaker, which operates in Lithuania, was chosen. Thus, a recommendation for those who follow this research would be to involve a higher number of bookmakers, and, therefore, look for arbitrage opportunities in the market. Finally, this research might suggest some ways of improvements for both bookmakers and punters. Analyzing the existing biases in the market, bookmakers can adjust their market prices in order to become more efficient and gain more profits from the bettors. Bettors, on the other hand, might develop and try various potential betting strategies.

6. REFERENCE LIST

Berkowitz J. P., Depken C. A., Gandar J. M. (2017). A Favorite-Longshot Bias in Fixed-Odds Betting Markets: Evidence from College Basketball and College Football. *The Quarterly Review of Economics and Finance*, 63, 233-239.

Constantinou, A. C. & Fenton, N. E. (2013). Profiting from Arbitrage and Odds biases of the European Football Gambling Market. *The Journal of Gambling Business and Economics*, 7(2), 41-70.

Croxson, K., & James Reade, J. (2013). Information and Efficiency: Goal Arrival in Soccer Betting. *The Economic Journal*, 124(575), 62–91.

Deschmaps, B. & Gerdaug, O. (2008). Efficiency in Horse Race Betting Markets: The Role of Professional Tipsters. Handbook of Sports and Lottery Markets. In D. B. Hausch & W. T. Ziemba (Eds.), *Handbook of Sports and Lottery Markets* (p.p. 341-354). North Holland.

Dowie, J. (1976). On the Efficiency and Equity of Betting Markets. *Economica*, 170, 139-150.

European Gaming and Betting Association. (2021). *Online Gambling Market (EU-28)*. Retrieved from https://www.egba.eu/eu-market

Fama F. E. (1970). Efficient Capital Markets: A Review of Theory and Empirical Work. *The Journal of Finance*, 25(2), 383-417.

Flepp R., Nuesch S., Franck E. (2017). The Liquidity Advantage of the Quote-driven Market: Evidence from the Betting Industry. *The Quarterly Review of Economics and Financ*e, 64, 306-317.

Forrest, D., & McHale, I. (2007). Anyone for tennis (betting)? *The European Journal of Finance*, 13(8), 751-768.

Franck, E., Verbeek, E., Nuesch, S. (2013). Inter–market Arbitrage in Sport Betting. *Economica*, 80(318), 300–325.

Goddard J. & Asimakopoulos I. (2004). Forecasting Football Results and the Efficiency of Fixed-odds Betting. *Journal of Forecasting*, 23, 51–6

Gray, P. K., & Gray, S. F. (1997). Testing Market Efficiency: Evidence from the NFL Sports Betting Market. *The Journal of Finance*, 52(4), 1725-1737.

Harkins, J. (2013). *Does NFL Spread Betting Obey the Efficient Market Hypothesis?* Retrieved from https://www.reed.edu/economics/parker/s13/312/fproj/JH.pdf

Kipkorir, J. W. (2017). *A Study of Behavioral Biases Present in Sports Betting Markets* [Bachelor's thesis, Strathmore University]. BBSE Research projects (2017). Retrieved from https://su-plus.strathmore.edu/bitstream/handle/11071/5377/A%20study%20of%20behavioral%20biases%20present%20in%20sports%20betting%20markets.pdf?sequence=1&isAllowed=y

Levitt, S. (2004). Why are Gambling Markets Organised So Differently from Financial Markets? *Economic Journal*, 114, 223–46.

Narayan, N. (2021). *Lithuania's Online Gaming Growth Fails to Offset Land-based Decline in 2020*. Retrieved from https://europeangaming.eu/portal/latest-news/2021/02/08/86219/lithuanias-online-gaming-growth-fails-to-offset-land-based-decline-in-2020/

Newall P. W. S. & Cortis, D. (2021). Are Sports Bettors Biased toward Longshots, Favorites, or Both? A Literature Review. *Risks*, 9(1), 1-9.

Ottaviani, M., & Sorensen, P. N. (2008). The Favorite-Longshot Bias: An Overview of the Main Explanations. Handbook of Sports and Lottery Markets. In D. B. Hausch & W. T. Ziemba (Eds.), *Handbook of Sports and Lottery Markets* (p.p. 83-101). North Holland.

Pachur, T. & Biele, G. (2007). Forecasting from Ignorance: The Use and Usefulness of Recognition in Lay Predictions of Sports Events. *Acta Psychologica*, 125(1), 99-116.

Paul, R. J. & Weinbach, A. P. (2014). Market Efficiency and Behavioral Biases in the WNBA Betting Market. *International Journal of Financial Studies*, 2, 193–202

Radiul A. & Janusonis L. (2020). *Betting Markets and Market Efficiency: Evidence from Higher Latvian Football League* [Unpublished bachelor's thesis]. Stockholm School of Economics in Riga.

Sauer, R. D. (1998). The Economics of Wagering Markets. *Journal of Economic Literature*, 36(4), 2021–64

Stanek, R. (2017). Home Bias in Sport Betting: Evidence from Czech Betting Market. *Judgment and Decision Making*, 12(2), 168-172.

Strong, N. & Xu X. (2003). Understanding the Equity Home Bias: Evidence from Survey Data. *The Review of Economics and Statistics*, 85, 307-312.

Teall, J. L. (2018). *Financial Trading and Investment*. Academic Press.

Thaler, R. H. & William. T. Ziemba. (1988). Anomalies-Parimutuel Betting Markets: Racetracks and Lotteries. *Journal of Economics Perspectives*, 2, 161-74.

Vaughan Williams, L. & Paton, D. (1997). Why is There a Favourite-longshot bias in British Racetrack Betting Markets? *Economic Journal*, 107(440), 150–158.

Vlastakis, N., Dotsis, G., Markellos, R. N. (2009). How Efficient is the European Football Betting Market? Evidence from Arbitrage and Trading Strategies. *Journal of Forecasting*, 28(5), 426–444.

Eurolygos lažybų rinka ir jos efektyvumas

Bakalauro darbas

Kiekybinė ekonomika

Vilniaus universitetas, ekonomikos ir verslo administravimo fakultetas

Vadovas – Milda Norkutė

Vilnius, 2021

Santrauka

40 puslapių, 8 lentelės, 4 iliustracijos, 27 šaltiniai.

Šiuo tyrimu siekiama įvertinti, ar Eurolygos lažybų rinka veikia efektyviai, įgyvendinant ekonometrinį modeliavimą ir patikrinant, ar nėra reikšmingų duomenų, kurie galėtų padėti prognozuoti sporto renginių rezultatus ir įvairius elgesio šališkumus rinkoje. Rašto darbas yra suskirstytas į keturis pagrindinius skyrius: literatūros apžvalga, metodika, pagrindinės išvados ir diskusija.

Literatūros apžvalgoje pateikiama teorinė informacija apie lažybų rinkas. Po to, apžvelgiama akademinė literatūra apie efektyvumą, regresijos modelius ir šališkumus. Pirmame skyriuje pateikiamos įvairios lažybų rūšys; antroje dalyje lažybų rinkos lyginamos su tipinėmis finansų rinkomis, siekiant ieškoti panašumų ir skirtumų; trečiajame skyriuje apžvelgiamas ekonometrinis modeliavimas, naudojamas kituose tyrimuose; ir, galiausiai, paskutiniame skyriuje pateikiami dažniausiai pasitaikantys elgesio šališkumai ir aptariamos kitų analitikų išvados.

Metodikos dalyje autorius pateikia analizei naudojamus duomenis ir metodus, naudojamus ekonometrinei daliai reikalingiems kintamiesiems apskaičiuoti. Taip pat, pateikiami įvairūs plėtojami ekonometriniai modeliai, kurie tikrina, ar yra kokių kintamųjų, galinčių padėti nuspėti įvykių baigtis. Jei tokių yra, lažybų rinka laikoma neveiksminga. Po to, šio tyrimo rezultatai ir išvados aptariami kitame skyriuje, kuriame EuroLeague lažybų rinka yra apibrėžiama kaip silpnos formos neefektyvia rinka.

Paskutiniame skyriuje apžvelgiamas atliktas darbas ir pateikiamos pagrindinės darbo išvados. Galiausiai, pateikiama šios analizės reikšmė ir rekomendacijos tiek lažybų organizatoriams, tiek lažybų dalyviams.

EUROLEAGUE BETTING MARKET AND ITS EFFICIENCY

Bachelor thesis

Quantitative Economics

Vilnius University, Faculty of Economics and Business Administration

Supervisor – Milda Norkutė

Vilnius, 2021

Summary

40 pages, 8 tables, 4 figures, 27 references.

This study aims to evaluate whether EuroLeague betting market is operating efficiently by implementing econometric modelling and checking for the existence of any significant data that might help to predict the outcomes of sports events and various behavioral biases in the market. The paper is divided into four main sections: the literature review, methodology, the main findings, and conclusion.

In the literature review, a theoretical background information on the betting markets is presented. Then, existing academic literature on the efficiency, regression models, and cognitive biases is reviewed. In the first section, various types of betting are presented; in the second part, betting markets are compared to the typical financial markets to look for similarities and differences; the third section reviews econometric modelling used in other researches; and, finally, the last section presents the most common behavioral biases and discusses the findings of other analysts.

In the methodology part, the author presents data used for the analysis and methods used to compute the variables needed for the econometric part. Also, various augmented econometric models are presented that check for significance of any observable variables. If such exist, the betting market is considered to be inefficient. After that, the results and findings of this research are discussed in the next section, in which the EuroLeague betting market is concluded to be weak-form inefficient.

The last section reviews the work done and concludes the main findings of the paper. Finally, the significance of this analysis and recommendations for both bookmakers and bettors are presented.