

DOI: [10.52950/ES.2021.10.1.004](https://doi.org/10.52950/ES.2021.10.1.004)

SPORTS MARKET AS A DATA SOURCE FOR ECONOMICS: WITH SPECIAL EMPHASIS ON BETTING AND FANTASY SPORTS

VOJTECH KOTRBA, ROBERT HOLMAN

Abstract:

Data collection for verification of economic models is often not easy. Data coming from laboratories can be distorted in comparison to data from the real environment. The sports environment offers an interesting source of data for research not only on the field of labour economics, thanks to its extensive records of performance, wages, and market prices of individual athletes. Sports betting and fantasy sports further extend this potential towards consumers. Data from this environment allow better recognition of the users' preferences, rational behaviour, or heuristics. This paper offers an outlining of the current state of economics focused on the use of the sports environment in research and shows some of the possibilities presented by sports betting and fantasy sports. As an example of that, the over-selection of athletes from the historically dominant teams in the fantasy league according to the Czech football competition in the season 2015-16 is investigated in the applied section. The results show the impact of this heuristic on the users' decision-making.

Keywords:

behavioural economics, fantasy sport, historically dominant teams, neoclassical economics, sports betting, data collection

JEL Classification: D01, L83, J24

Authors:

VOJTECH KOTRBA, Prague University of Economics and Business, Czech Republic, Email: vojta.kotrba@post.cz

ROBERT HOLMAN, Prague University of Economics and Business, Czech Republic, Email: robert.holman@vse.cz

Citation:

VOJTECH KOTRBA, ROBERT HOLMAN (2021). SPORTS MARKET AS A DATA SOURCE FOR ECONOMICS: WITH SPECIAL EMPHASIS ON BETTING AND FANTASY SPORTS. International Journal of Economic Sciences, Vol. X(1), pp. 53-70., 10.52950/ES.2021.10.1.004

Supported by: the grants No. VŠE IGA 506020 (IGS F5/4/2020) and VŠE IGA 506010 (IGS F5/2/2020) of the Internal Grant Agency by Faculty of Economics, Prague University of Economics and Business.

Acknowledgements: The authors would like to express his gratitude to Kateřina Kotrbová and Pavel

Kotrba and Anežka Kotrbová for their help and consultation. Thanks also include Seznam inc. for providing the fantasy league raw data.

1 Introduction

Sports fandom, these days, is not just based on visited matches, but also on many affiliate products such as sports betting or fantasy sports. They offer entertainment for users and, moreover, an excellent area for research, as these technologies allow uninterrupted data collection.

Economic research using sports betting and fantasy sports is a marginal area in economics. The use of sports betting is certainly a larger area, but the possibilities for its application can be limited. The relatively new fantasy sports environment allows for wider use thanks to its greater interactivity. Although the number of papers using this environment is rising, many editors and researchers still need to be convinced of the appropriateness of this data source.

Data is essential for the social sciences to continue their development and ensure that their findings are up to date. The testing of new ideas in the laboratory environment is currently widespread as it allows relatively cheap and rapid conclusions. However, basing findings only on laboratory data could lead to errors (Kagel and Roth, 2016). People may behave differently if they are supervised and their motivation may vary as well. It is therefore important to confirm the theoretical models on data outside the laboratory to confirm economic models.

Obtaining appropriate data for social science research may not be easy. In many areas, data is owned by private entities that may not be willing to share it with researchers. In other areas, there is simply no data, since no one records or stores it. Therefore, the sports environment has many benefits such as athletes' performance records, availability of their salaries, audience interest records, and many others (Rosen and Sanderson, 2001). This leads to the frequent use of the sports environment for economic research (Kahn, 2000) or social science research in general (Gratton and Jones, 2010).

In addition to the research on the sports labour market, it is currently possible to examine the preferences of sports spectators and fans in more detail. This is done by using tools that facilitate the collection of data in the sports environment in a natural way. The traditional demand of fans can be examined using viewership at stadiums (Lahvička, 2010) or on television (Sung et al., 2019). The areas such as betting (Thaler and Ziemba, 1988) or fantasy sports (Dwyer and Drayer, 2010) could be also used. These areas can be an interesting environment for a neoclassical approach based on rationality, as both betting and fantasy sports are about winning through optimal strategy. The potential for behavioural research is to detect imperfections and deviations from the optimal strategy. Given the high amount of time required by betting and fantasy sports, players may not be willing to actually seek the optimal strategy, and it is, therefore, possible to detect the heuristics they use.

Contemporary neoclassical economics has an extensive background in which many theoretical models can be found to be tested and validated. On the other hand, behavioural economics is a developing area, where new theoretical models are still emerging, but existing models also need to be tested. At the same time observations from both real and laboratory environments can help discover new explanations of social phenomena and create new models.

The following paper aims to outline the current state of the economy using the sports environment and to show the possibilities that sports betting and fantasy sports offer. Anchoring this area in research can be seen in the decision theory, rational choice or behavioural approach. Blending could also occur, for example, with the game theory or the behavioural game theory.

In order to demonstrate the value of the use of the new sports-based data, the over-selection of athletes from the historically dominant teams is tested in the applied part on data from the fantasy league according to the highest Czech football competition in the season 2015-16. The results show that the users are indeed influenced by this heuristic, but the performance of football players has the greatest influence. Nevertheless, the results of this section are not and should not be the main added value of this paper, which is rather the above-mentioned overview and presentation of the sports environment for applied research in economics.

The paper is structured as follows. In Section 2, the issue of data acquisition for economic models is discussed. Consequently, in Section 3, the advantages of sports environment usage are emphasized. The following Section 4 and Section 5 show the potential of using sports betting and especially fantasy sports for data acquisition and research in economics and possibly other social sciences. In Section 6, the presence of the over-selection of athletes from the historically dominant teams in the Czech fantasy league is tested. Possible complications that the researchers might encounter in this area are presented in the conclusion which is Section 7.

2 Data acquisition for applied economics

Important challenge economists face is the collection of data. They need it to test the functionality of their models, both neoclassical and behavioural. In principle, the origin of the data can be divided into two groups.

The first group is data coming from situations where respondents realize that data about them is being collected which will be used for research. A specific example could be a questionnaire survey (Marsden and Wright, 2010), either using the Internet or other elements of computer technology, as well as field research. Experiments in an economic laboratory may be another example (Smith, 1976). At present, this approach is very popular in many social sciences, such as economics, psychology or sociology. Nevertheless, laboratory results bring some risks. Often a researcher cannot be sure that the experimental conditions fully correspond to reality. For example, survey participants could behave differently in laboratory conditions than in reality. In economic experiments, scientists often try to motivate participants with real profit opportunities to bring their behaviour closer to reality.

A big advantage of laboratory experiments is a relatively easy way to obtain data. Assuming a correct design of the experiment, getting the answers from the respondents is only a matter of a sufficient number of people who arrive at the laboratory, which financial reward often allows.

The second group is data from the natural environment. People in the real world were motivated in particular situations and had no idea that anyone would use their behaviour records for research. This data is called field data. A special type is a natural experiment where a big change occurs in the real environment and the situation before and after the change is investigated (List, 2007).

A disadvantage of natural experiments and field data is the great difficulty of their acquisition. At present, human behaviour and decision-making are often recorded, but certainly not in all areas of economic research interest. And once human behaviour was recorded and converted to usable data, then it is usually not publicly available. Such data is often owned by private companies that want to use it for their own profit rather than for scientific research. Nevertheless, there may be negotiations between scientists and private companies to provide at least anonymized data for

scientific purposes (Einav and Levin, 2014). Data from the state that the public sector decides to publish may be an alternative.

In reality, it is often quite difficult to take the behaviour of all the subjects into account to whom we want to draw conclusions. Therefore, there is often so-called sampling, where research is carried out only on some of the subjects and not on the whole (Babbie, 2015). However, it is important for valid conclusions that the selected sample is representative. This means that it corresponds to the whole, only in a reduced proportion. For example, if the total is half women and half men, it is not possible for the selected sample to consist of 75 per cent men and only 25 per cent of women. The ratio of men and women should remain the same (at least approximately) as on the whole.

3 Sports as a suitable environment for research

In this section, there are shown advantages of using sports environment to investigate social science questions, specifically applied economics.

The sports environment has many interesting advantages for economic research. Sports can be seen as an entertainment industry or leisure activity. In terms of leisure activities, many people spend their free time playing sports (Izquierdo Sanchez et al., 2016). They, therefore, prefer it over other consumption options offered to them. At the same time, they often invest a considerable part of their income into sports. However, much of the economic research is focused more on sports as an entertainment industry, where the consumption of fans and the possibility of managing this consumption through marketing can be researched. Millions of people follow sports in newspapers and magazines (Eagleman, 2011), on television (Sung et al., 2019), on the Internet (Hur et al., 2007) or directly at stadiums (Lahvička, 2010). Apart from the film industry and computer games (where e-sport already plays a big role), sport is undoubtedly one of the most used entertainment environments. In North America alone, revenues from sports markets reached \$ 69.14 billion in 2017 (Statista, 2019).

One of the most important economic areas where the sports environment is used for research is the labour market. In sports, there are detailed records of athletes' performance, which are also kept for a long time and often publicly available (Rosen and Sanderson, 2001). Data availability is one of the greatest advantages of the sports environment (Schumaker et al., 2010). In many other sectors, it is often not possible to measure the performance of individual workers. Since performance is not recorded elsewhere, there is no data to determine the relation of performance and wages paid to workers. Thus, the sports environment is interesting for wage discrimination research.

Sports labour markets have been appearing in social sciences papers since the middle of the last century (Rottenberg, 1956). An important milestone is a possibility of having more foreigners in the team and the introduction of the Bosman rule, which gives athletes the opportunity to transfer free of charge after their contract expires. This makes the sports labour market very close to the regular labour market (Frick, 2007). There are, of course, specifics regarding the organization of the sports labour market. A sector with a small number of firms and a large number of workers should show monopsony rents (Kahn, 2000), but this only happens in rigid markets where stricter regulations apply, such as salary ceilings or the complications in transfers of athletes between clubs.

Athletes' wages depend primarily on factors such as age and experience (MacDonald and Reynolds, 1994), the number of goals scored, the players' position (midfielders earn the most, goalies earn the least) or its fame (Google searches) (Frick, 2007). Superstar players receive higher wages than others, and the viewer demand for them is increased (MacDonald and Reynolds, 1994; Hausman and Leonard, 1997; Frick, 2007; Kotrba, 2019). However, they bring additional revenues to the teams (merchandising, sponsors). Their presence in the team can also be perceived as a positive externality for the teams they face (Hausman and Leonard, 1997). In addition, the presence of a star player increases attendance at matches and TV viewing.

Also, racial discrimination has been one of the main issues in a team sport for a long time (Scully, 1973). In the 1950s and 1970s, discrimination against blacks was proved. However, this phenomenon subsided in some leagues in the 1990s (Kahn, 2000) as teams with a higher number of blacks achieve statistically better results at the same labour cost (Szymanski, 2000). Nevertheless, preferences for whites can be still found among fans (Kahn, 2000).

4 Sports betting

Another industry interesting for economic research is sports betting. Factors like maximization of utility, uncertainty or expected income play a role in betting, which are factors frequently used in economic theory. Moreover, this environment also produces data suitable for subsequent research, which is often the case (Williams, 1999).

The amount of money bet in Europe in 2017 exceeded 94.7 billion Euro (European Gaming and Betting Association, 2018). In the USA, the overall spending on the hazard in 2017 reached 109.0 billion dollars (American Gaming Association, 2018).

The beginnings of sports betting reach back to Great Britain, where horse race betting was invented (Huggins, 2007) to be later joined by other sports like football or cricket (Forrest and Simmons, 2003). Recently, online betting has become really popular (Lopez-Gonzalez et al., 2017). In the USA, however, the situation was different compared to benevolent Europe (Reiche, 2013). Before 2018 a law called the Professional and Amateur Sports Protection Act was in effect, which made all sports betting illegal. The only exception from this law was the state of Nevada. This situation resulted in widespread illegal betting all around the country (Munting, 1996). Since 2018, sports betting can be legalized by other states, which is also happening.

The first area using sports betting is the theory of effective markets (Figlewski, 1979). This theory assumes that traders change their behaviour only on the basis of objective information with a real impact on the market itself. In the case of sports betting, this would mean that the betting odds are set so that the bettors and bookmakers as a whole end up on zero. Eventually, overhead costs may be added for betting providers as payment for consumption. It would not be possible to find overvalued or undervalued odds on the market. Thus, this theory implies, that the success rate of betting cannot be improved by historic or current data analysis. Nevertheless, research in this field points out that with the right strategy and analysis, it is, in fact, possible to make profits (Graham and Stott, 2008).

Another area using sports betting is the research of bettors' rationality. If the assumptions of neoclassical economics were applied, bettors would place only rational bets and make optimal decisions, which is not the case in reality (Thaler and Ziemba, 1988). Among bettors, only a fraction can be found who try to exploit the available information for their benefit (Asch et al.,

1982). The later betting individuals should have better information, which is, in reality, happening only in some markets (Gramm and McKinney, 2009).

Aside from the economic benefits for betting offices and the consumption of bettors, it is also necessary to mention the negative impacts betting may bring. For the majority of bettors, betting is only an innocent entertainment, however, there is a group of consumers who might develop an addiction to it (Winters et al., 1995).

Also, the influence of sports results by athletes for money might be connected to the betting industry (Wolfers, 2006). Corruption scandals like that subsequently shed a bad light on the whole sports industry.

The field of betting offers further directions of research focused on the intersection with behavioural economics and the comparison with the approach of neoclassical economics. Also, betting on individual members of a sports team may be useful to the research of bettors' preferences and heuristics.

5 Fantasy sport

Fantasy sport is an online activity, where its users play the part of sports managers and build their own squads within a limited budget. The athletes who are chosen by the users bring points for their real-life performance. Thus, the course of fantasy leagues follows the course of the real competitions. The choice of individual athletes allows gathering records of the users' behaviour in great detail and expands the possibilities of economic research.

For example, if a user took part in the fantasy league that was used in the applied part of this paper, he or she was given a virtual budget of 100 million after the registration on the website, to buy football players for his or her team. The user had to always buy one goalkeeper, 3 to 5 defenders, 3 to 5 midfielders and 1 to 3 forwards to reach the overall number of 11 players. Among those football players, the user had to mark a captain who brought double the points. Each athlete had a price assigned by the fantasy league authors, which was influenced by factors such as the player's performance, fame or field position. Subsequently, the user gained points according to his or her football players' performance in the given round. This process was repeated before each round. The user's logical goal was to gain as many points as possible (Dwyer et al., 2011).

The interest in fantasy sports is keeps growing, which is pointed out by the data of Fantasy Sports and Gaming Association. In 2017, fantasy sports were played by more than 59.3 million Americans and Canadians and this number has increased even more since then (Fantasy Sports and Gaming Association, 2019).

While in Northern America, sports like American football, baseball, basketball, and ice hockey are popular, the most popular sport in Europe is football, by far. The official fantasy league according to the most famous football league, Premier League, which is played in England; was played by more than 6.3 million registered users in season 2018/19 (Fantasy Premier League, 2019). In comparison to 2013, when the game was played by a mere 3.5 million users (Bryson and Chevalier, 2015), the number rose by 80 per cent.

Harvard professor William Gamson is considered the founder of fantasy sports (Burton et al., 2013). In 1960 he developed a game based on real athletes' drafts into the virtual teams of the individual users, who then received points according to their athletes' real performance. Given the

time the game was run, traditional mail was used as means of communication between the participants. The users had the competence to choose athletes, alternate them between the bench and the field, sell the players and buy new ones. That with the goal of gaining as many points as possible.

This concept spread quickly into other sports and as soon as 1963 rules for competition according to American football was set and Greater Oakland Professional Pigskin Prognosticators League was founded, which, to this day, still works (Burton et al., 2013). It must be noted that it still keeps the original way of drafting with the use of pen and paper. The next important milestone is the creation of the Rotisserie Baseball System by journalist Daniel Okrent. The paper on this game system spread quickly across the United States and led to the foundation of many local leagues (Schwarz, 2004). Another step was taken in 1989 when the Dugout Derby company made it possible to play fantasy sports based on Rotisserie Baseball System over the phone (Burton et al., 2013). This product met such great success, that the offer was extended to NFL, PGA, and NASCAR as well.

Nevertheless, the biggest boom of fantasy sports came only with the expansion of the Internet. In 1999, Fantasy Sports and Gaming Association was founded (the original name being Fantasy Sport Trade Association), which monitors the development of this market and issues interim reports on the industry. FanDuel and DraftKings are the biggest operators in the USA which run mostly daily fantasy sports.

While the traditional fantasy sports is in progress throughout the whole season of the given league (NFL, NHL, MLB, Premier League, and more), the daily version could be considered a speeded up format. In the morning the users set up their squads and in the evening the game is already decided.

This variation of fantasy sports has become truly popular in North America. Traditional fantasy sports is often played by a group of friends who measure their strategies and abilities. Daily fantasy sports differ. The users often pay an entrance fee in the morning to cash the collected amount in the evening in case they win. Minus the due overheads for the respective league operators, that is. This form of fantasy sports leans more towards gambling, but research points out that other than luck, correct strategies and abilities are essential (Evans et al., 2018).

As mentioned before, sports betting was illegal in the USA for a long time, with the exception of Nevada (Professional and Amateur Sports Protection Act). In those times, fantasy sports stayed in the grey zone between legality and illegality. In order for a game to be considered a lottery, it has to meet three characteristics according to American law. Those are entrance fees, prize, and chance (Moorman, 2008). In the case of fantasy sports, the main question is whether the game is based on chance or abilities. Many papers focused on this topic in the USA (Holleman, 2006; Boswell, 2008). It was up to individual states' courts to resolve the disputes on legality. Most of the states allowed fantasy sports as a game based on abilities, with the single exception of Nevada. Ironically, with Nevada being the only state allowing gambling, it only prompted the operators to get a license for betting activities.

Another attempt of setting fantasy sports outside the law came with the dispute over the ownership of statistical data from sports used in fantasy leagues. Sports leagues claimed those data and demanded a fee for their use. The controversy on the subject is investigated by some papers of American authors such as Massari (2006) or Weaver (2010). However, the court ruled that basic characteristics and statistics can be used in fantasy sports without charge.

A great task for sociologists, economists, but particularly sports marketers and managers was to find out how fantasy sports influence traditional fandom and the sports industry as a whole. Primarily, it was necessary to determine who are fantasy sports users and if there is an overlap with traditional fans.

Surveys show that fantasy sports users are mostly men (Davis and Duncan, 2006), often enthusiastic sports fans (Bernhard and Eade, 2005). It was also proved that traditional fandom influences fantasy sports as dedicated fans demand athletes from their favourite club more (Dwyer, 2011). However, many users prefer the victory of their own fantasy squad over the victory of their favourite club (Lee et al., 2013). And conversely, fantasy sports influences traditional sports watching in a clearly positive way. It increases viewership in TV as well as attendance at stadiums (Nesbit and King, 2010a, 2010b).

In economics, data from fantasy sports are increasingly used for determining consumers' preferences (Bryson and Chevalier, 2015) or to verify the rationality of their behaviour and the possible use of heuristics (Kotrba, 2019). Nevertheless, the potential for further research is vast. It would be possible to better explore the areas that have been already used in some papers, such as racial discrimination (Bryson and Chevalier, 2015) or the superstar effect (Kotrba, 2019). Another advancement could be seen in areas such as the economics of beauty (Mobius and Rosenblat, 2006), attitude towards foreigners on sports (Wilson and Ying, 2003) or in uncovering of the strategies used by fantasy sports users (Becker and Sun, 2016).

Sections 2-5, just presented, show that applied economics must seek appropriate natural environments for data collection in order to validate theoretical models. The sports environment proves to be very suitable, in particular, because of the existence of detailed records of athletes' performances, wages and prices. Similarly, there are precise records of fan behaviour. Two interesting environments for investigating fan behaviour appear to be sports betting and fantasy sports. Sports betting is a traditional tool for investigating fan preferences and rationality in the economics literature. Fantasy sport is only just coming to the fore, but it provides many additional benefits. First of all, it allows fans' preferences to be individualized to individual athletes. It, therefore, allows more precise targeting of economic research results.

In order to support these theoretical conclusions, the following section 6 presents a specific application of data from fantasy sport to investigate fan behaviour. This section thus aims to demonstrate possible applications in applied economics. However, this is only an example and is not intended to outweigh the value of the overall review of the issue of the use of sports betting and fantasy sports data in social sciences.

6 Historically dominant teams

This empirical section presents an example of the use of data from a fantasy sports environment. Specifically, a bias that shows that fans will overestimate the performance of the historically dominant teams is investigated. The theoretical anchoring of this bias can be explained in several ways.

The first possible explanation could be the following; historically successful teams tend to have more fans because a team's victory attracts more fans who would like to share the feeling of victory. This expands the ranks of traditional loyal fans who cheer for the team even when the

team is not doing well (Hornby, 2005). Fans of clubs, then, will logically have a greater urge to include their favourite team's athletes on their fantasy team (Kotrba, 2019).

Another possible explanation may be an incorrect estimate of the probability that high-performing teams will continue to win. An anchor for this claim can be found in the representativeness heuristic (Tversky and Kahneman, 1974). This heuristic is used as a help in uncertain decision-making and describes cognitive distortion where an individual makes his or her decision based on the similarity of the situation with a so-called representative situation.

An example from the sports environment is a constantly winning team. In case the victorious row lasts long enough, an individual might automatically assume that it will win the next match as well. In situations like this, so-called regression to the mean often appears (Tversky and Kahneman, 1974). It is a phenomenon where an extreme value tends to return to normal. However, an individual often expects the value to remain extreme. Therefore, it is probable that a constantly winning team will eventually lose, despite the individual's expectation. This false assumption can subsequently lead to mistakes (biases) in behaviour in various environments.

The heuristic of representativeness has been already tested in the betting environment where it should appear in overestimating the probability of victory of the team or athlete who had been previously successful. This false perception of probability might affect both bettors and betting offices. However, Tassoni (1996) shows data from betting on American football between 1976 and 1979 that the heuristic of representativeness almost does not appear. On the contrary, Woodland and Woodland (2015) detect the presence of the heuristic on later data from the years 1998 to 2010, at least in a weak form.

However, it must be stressed that there are certainly other explanations for the over-selection of athletes from the winning team into fantasy squads. In some leagues, some teams may outperform others in terms of revenue and make purchases of the best athletes to maintain their above-average performance. Or they can have better managers or youth academy.

6.1 Data and hypothesized model

The used dataset is in the form of panel data where the rounds of fantasy league act as a time period where the users' demand for individual football players is monitored. Thus, the dependent variable aims to quantify this demand. Considering the fact that a different number of players took part in the single rounds, it is defined as a percentage of squads into which an athlete was chosen. For example, in case 8,000 users would enter a round and 2,000 of them would choose a particular football player, the dependent variable will have a value of 0.25 in this observation. In the model, this variable is used in a logarithmic form following the previous literature (Kotrba, 2019; 2020; 2021) and skewness and it is marked l_demand .

Independent variables aim to show what is the motivation of the users to include a football player in their squads. The player's performance should play a lead role. There are three types of performance included in the model in accordance with the previous literature (Bryson and Chevalier, 2015; Kotrba, 2019; 2020; 2021). Long-term performance represents an athlete's stability in scoring. The variable is marked in the models as p_long . Short-term performance represents the athlete's form prior to the observed match and it is marked p_short in the models. The current performance aims to filter out the information that is available shortly before the match. Those are for example card penalties, injuries or comments of the teams' coaches on the players' inclusion in the squad. This variable is marked $p_current$.

Considering the fact that players on different field positions gain points differently, the dummy variable for game positions are included in the model. The variables are defender, midfielder and forward. The variable for a goalkeeper is omitted in order to avoid perfect multicollinearity.

The main variables are dummy variables for athletes from the clubs which were successful in the season preceding the monitored one. Particularly, the first four teams of the previous season are monitored, which are those that entered European competitions in the monitored season. Those are FC Viktoria Plzeň (viktoria), AC Sparta Praha (sparta), FK Jablonec (jablonec) and FK Mladá Boleslav (boleslav). In brackets, the markings in the model are stated. Thus, dummy variables divide the whole dataset into observations of the football players from the four most successful clubs of the previous season and observations of the athletes from the clubs below the fourth place. The heuristic of representativeness will be present if the players from the four best teams will be chosen into the squads more often than the others. That is in case the coefficients for those clubs will be statistically significant in the model with a positive influence. Belonging to a club should be an irrelevant factor as the only relevant factor should be the athletes' performance (Kotrba, 2019; 2020; 2021).

The biggest disadvantage of the dataset is a missing variable representing the price of football players. However, this information is not available in the dataset, so it cannot be included. Nevertheless, it can be expected that the fantasy league authors tried to set the price so that the individual users' squads differed in order for the competition to be interesting. The users could not buy all the best scoring players as they would be too expensive. Thus, the price will with a high probability correspond with performance, which is sufficiently represented in the model.

6.2 Econometrics models and results

The data used in this paper to test the over-selection of athletes from the historically dominant teams comes from a fantasy league operated by a company Seznam inc. in season 2015-2016 played according to the Czech highest football competition. The users' squads from the 17th to 27th round are at disposal. A total number of 8,036 users took part in the game who filled 53,951 squads into which they chose 482 different football players. Based on this data, an econometric model is assembled.

Many footballers from the full list were not chosen in some of the rounds. Therefore, the analysed data present an unbalanced panel. The results of the Hausman test applied to the intended model (m -value = 201.01, p -value < 0.001) lead to rejecting the random-effects model. Furthermore, some characteristics that may affect the demand for footballers are time-invariant which makes the use of fixed effects impossible, however, the results of the F test for no fixed effects applied to the intended model (F -value = 5.22, p -value < 0.001) lead not to rejecting the fixed effects model.

To avoid correlation between observations of an individual athlete in different rounds, two possible approaches were chosen. The first approach was to create a separate model for each round, which means eleven models in total. These models were treated as cross-sectional data using the ordinary least-squares method. The second method was to cluster the athletes and create one model for the entire observation period. Again, the ordinary least-squares method was used.

For both approaches, it was necessary to choose different expressions of the athlete's performance. In the first approach, performance was constructed as follows. Long-term performance represents the average points from the previous half of the season, ten and five last

matches. The reason is the users' gradual forgetting (Bryson and Chevalier, 2015). Short term performance shows the number of points gained in the round previous to the monitored. Current performance represents the number of points gained in the monitored round.

In the second approach, performance was constructed as follows. Long-term performance represents the points from the previous season. Short term performance shows the sum of points gained in the rounds previous to the monitored period in the season 2015-2016. Current performance represents the number of points gained in the monitored period. To avoid confusion, the variables are marked as follows p_long2, p_short2 and p_current2.

The results of the first approach can be found in Table 1 and the results of the second approach can be found in Table 2.

Table 1 Econometrics models for separate rounds (dependent variable: L_{demand}).

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
const.	-6.143***	-8.052***	-7.595***	-8.066***	-8.2***	-8.123***	-7.965***	-8.192***	-7.86***	-7.651***	-7.352***
p_current	0.229***	0.233***	0.211***	0.281***	0.252***	0.161***	0.234***	0.258***	0.28***	0.225***	0.253***
p_short	0.032	0.154***	0.233***	0.063**	0.185***	0.101**	-0.027	0.053	0.095***	0.114***	0.042***
p_long	0.378***	0.75***	0.881***	0.883***	0.787***	0.894***	1.018***	0.988***	0.972***	0.964***	1.024***
defender	0.244	0.103	-0.17	0.095	0.197	0.224	0.044	0.237	0.245	0.081	-0.117
midfielder	-0.251	-0.349	-0.744***	-0.532**	-0.302	-0.077	-0.458*	-0.435*	-0.672***	-0.885***	-1.156***
forward	-0.049	-0.264	-0.578*	-0.445	0.141	0.327	-0.169	0.124	-0.124	-0.24	-0.463*
viktoria	1.389***	1.120***	-0.17	0.337	0.395	0.897***	-0.273	1.061***	-0.028	0.29	-0.868***
sparta	1.211***	1.271***	0.491	1.691***	1.31***	0.853**	0.606*	0.967***	0.71**	1.054***	1.107***
jablonec	-1.128***	1.975***	-0.224	1.541***	0.308	0.505	1.254***	0.394	1.544***	-0.31	1.091***
boleslav	0.093	0.969***	1.215***	1.038***	1.213***	0.343	1.992***	1.601***	0.294	0.529	0.485
Adj. R²	0.411	0.618	0.608	0.682	0.722	0.6	0.635	0.667	0.688	0.635	0.648
p-value	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
(F-test)											
Number of obs.	242	349	365	358	346	335	340	325	361	339	371

Note: Significance at 10 per cent is marked *; at 5 per cent is marked **; at 1 per cent is marked ***

Table 2 Econometrics model for the whole observed period (dependent variable: I_demand).

	(12)
const.	-8.494***
p_current2	0.007***
p_short2	0.026***
p_long2	0.120***
defender	-0.078
midfielder	-1.196***
forward	-0.708***
viktoria	-0.172
sparta	0.944***
jablonec	1.026***
boleslav	0.907***
Adj. R²	0.86
p-value (F-test)	<0.001
Number of obs.	482

Note: Significance at 10 per cent is marked *; at 5 per cent is marked **; at 1 per cent is marked ***

The adjusted R² shows that all models for separate rounds explain about 60 per cent of variation explained by only the explanatory variables that actually impact the dependent variable except for the first model. The resulting coefficients for performance variables show that users primarily monitor the performance of football players. Long-term and current performances are statistically significant at 1 per cent in all models and short-term performance is not significant even at 10 per cent only in three models. In model 12, all three variables are significant at 1 per cent. Long-term performance has the greatest impact.

Game position variables are included only to eliminate different points scoring.

The significance and direction of the variables of the most successful clubs show that the football players of these clubs are actually chosen more in the squads. However, for AC Sparta Praha, FC Viktoria Plzeň and Jablonec, it is also possible to find rounds with negative influence. This is true only for one round for each club. While the positive influence occurs ten times for AC Sparta Praha, four times for FC Viktoria Plzeň and five times for FK Jablonec. FK Mladá Boleslav has six positive influences and no negative. Model 12 show overall positive influence of AC Sparta Praha, FK Jablonec and FK Mladá Boleslav.

A possible explanation for the lower significance of FC Viktoria Plzeň athletes is their price, in case it was set too high. The fantasy league authors might attempt to lower the number of this club's football players in the squads as they might have expected their high performance. This

effect is not discovered with the other teams, but it could be present with the overall winner of the previous season.

Considering the confirmation of significance in the case of three out of four clubs in model 12; it is possible to attribute influence to the historically dominant teams. Therefore, fantasy sport is an environment where this heuristic is in effect and influences the behaviour of its users. This causes the deviation from the optimal strategy that should be based primarily on the performance of athletes. Nevertheless, performance remains the key factor in choosing football players.

7 Conclusion

This paper aims to highlight the advantages that the sports environment offers to economic research. A special focus is on the use of sports betting and fantasy sports, which further develop those advantages. Thanks to those areas being operated online, data emerge suitable for utilization in research that allows to better capture fans' preferences and behaviour. The available data can enrich even other social sciences.

In particular, the presence of the over-selection of athletes from the historically dominant teams is being tested in this paper in a fantasy league played according to the Czech highest football competition in season 2015-2016. The results suggest that the historically dominant teams indeed plays a role in the users' decision-making process, even though the main influence is still performance.

Neither the use of data from sports betting nor fantasy sports are completely without trouble. One of the possible complications which are shared with many other industries is the non-public access to the data. They are usually owned by a private company that does not have to be willing to share them. However, in the case of sports betting and fantasy sports the data for sure exist, which is an advantage in comparison to other industries, where the data do not necessarily exist. Thus, it is only a question of negotiation. A good opportunity for obtaining data is the planned discontinuation of a specific fantasy league, which means the company will not use the collected data in the future, so it may be willing to give them up.

Another problem might lay in the identification of the group to which the conclusions from the research can be drawn. Sports bettors are a specific, but also large enough group (American Gaming Association, 2018; European Gaming and Betting Association, 2018). In the case of fantasy sports, the situation is slightly more complicated. Nevertheless, previous papers show that the group of fantasy sports users overlaps to a great extent with sports fans in general (Bernhard and Eade, 2005). Even though the group of fantasy sports users does not necessarily have to be big, which is rather the case of Europe (Fantasy Premier League, 2019) than North America (Fantasy Sports and Gaming Association, 2019); the group of fans, in general, is a sample sufficient in every way, so that it makes sense to deal with it.

Overall, it is necessary to admit that economic research using sports betting and fantasy sports environment is only a complement aiming to refine the conclusions from other areas. Nevertheless, quality data from the real environment render those environments interesting for economic research. It is true that to make macroeconomic conclusions based on games from the sports environment may not be the best way in many aspects. However, in the research of consumer behaviour, those areas offer great potential for enrichment of the current state of knowledge both from the viewpoint of neoclassical and behavioural economics.

8 References

- American Gaming Association. (2018). *Economic Impact of the US Gaming Industry*. Retrieved from <https://www.americangaming.org/wp-content/uploads/2018/06/OE-AGA-Economic-Impact-US-2018-June.pdf>
- Asch, P., Malkiel, B. G., and Quandt, R. E. (1982). Racetrack betting and informed behavior. *Journal of financial economics*, 10(2), pp. 187-194. DOI: 10.1016/0304-405X(82)90012-5
- Babbie, E. R. (2015). *The practice of social research*. Nelson Education.
- Becker, A., and Sun, X. A. (2016). An analytical approach for fantasy football draft and lineup management. *Journal of Quantitative Analysis in Sports*, 12(1), pp. 17-30. DOI: 10.1515/jqas-2013-0009
- Bernhard, B., and Eade, V. (2005). Gambling in a fantasy world: An exploratory study of rotisserie baseball games. *UNLV Gaming Research and Review Journal*, 9(1), pp. 29-42.
- Boswell, J. (2008). Fantasy Sports: A Game Of Skill That Is Implicitly Legal Under State Law, And Now Explicitly Legal Under Federal Law. *Cardozo Arts and Entertainment Law Journal*, 25(3), pp. 1257-1278.
- Bryson, A., and Chevalier, A. (2015). Is there a taste for racial discrimination amongst employers? *Labour Economics*, 34, pp. 51-63. DOI: 10.1016/j.labeco.2015.03.002
- Burton, R., Hall, K., and Paul, R. (2013). The historical development and marketing of fantasy sports leagues. *The Journal of Sport*, 2(2), pp. 6. DOI: 10.21038/sprt.2013.0225
- Davis, N. W., and Duncan, M. C. (2006). Sports knowledge is power: Reinforcing masculine privilege through fantasy sport league participation. *Journal of Sport and Social Issues*, 30(3), pp. 244-264. DOI: 10.1177/0193723506290324
- Dwyer, B. (2011). Divided loyalty? An analysis of fantasy football involvement and fan loyalty to individual National Football League (NFL) teams. *Journal of Sport Management*, 25(5), pp. 445-457. DOI: 10.1123/jsm.25.5.445
- Dwyer, B., and Drayer, J. (2010). Fantasy sport consumer segmentation: An investigation into the differing consumption modes of fantasy football participants. *Sport Marketing Quarterly*, 19, pp. 207-216.
- Dwyer, B., Shapiro, S. and Drayer, J. (2011). Segmenting Motivation: An Analysis of Fantasy Baseball Motives and Mediated Sport Consumption. *Sport Marketing Quarterly*, 20(3), pp. 129-137.
- Eagleman, A. M. (2011). Stereotypes of race and nationality: A qualitative analysis of sport magazine coverage of MLB players. *Journal of Sport Management*, 25(2), pp. 156-168. DOI: 10.1123/jsm.25.2.156
- European Gaming and Betting Association. (2018). EU Market. Retrieved from <https://www.egba.eu/eu-market/>
- Einav, L., and Levin, J. (2014). Economics in the age of big data. *Science*, 346(6210), pp. 1243089. DOI: 10.1126/science.1243089
- Evans, B. A., Roush, J., Pitts, J. D., and Hornby, A. (2018). Evidence of Skill and Strategy in Daily Fantasy Basketball. *Journal of Gambling Studies* 34(3), pp. 1-15. DOI: 10.1007/s10899-018-9766-y
- Fantasy Premier League. (2019). Retrieved from <https://fantasy.premierleague.com/>

- Fantasy Sports and Gaming Association. (2019). *Media kit*. Retrieved from <https://thefsga.org/wp-content/uploads/2019/04/2018-19-Press-Kit.pdf>
- Figlewski, S. (1979). Subjective information and market efficiency in a betting market. *Journal of Political Economy*, 87(1), pp. 75-88. DOI: 10.1086/260740
- Forrest, D., and Simmons, R. (2003). Sport and gambling. *Oxford Review of Economic Policy*, 19(4), pp. 598-611. DOI: 10.1093/oxrep/19.4.598
- Frick, B. (2007). The football players' labor market: Empirical evidence from the major European leagues. *Scottish Journal of Political Economy*, 54(3), pp. 422-446. DOI: 10.1111/j.1467-9485.2007.00423.x
- Graham, I., and Stott, H. (2008). Predicting bookmaker odds and efficiency for UK football. *Applied Economics*, 40(1), pp. 99-109. DOI: 10.1080/00036840701728799
- Gramm, M., and McKinney, C. N. (2009). The effect of late money on betting market efficiency. *Applied Economics Letters*, 16(4), pp. 369-372. DOI: 10.1080/13504850601018577
- Gratton, C., and Jones, I. (2010). *Research methods for sports studies*. Routledge.
- Hausman, J. A., and Leonard, G. K. (1997). Superstars in the National Basketball Association: Economic value and policy. *Journal of Labor Economics*, 15(4), pp. 586-624. DOI: 10.1086/209839
- Holleman, M. C. (2006). Fantasy football: Illegal gambling or legal game of skill. *NCJL and Tech.*, 8, pp. 59-80.
- Hornby, N. (2005). *Fever pitch*. Penguin UK.
- Huggins, M. (2007). Betting, sport and the British, 1918-1939. *Journal of Social History* 41(2), pp. 283-306. DOI:10.1353/jsh.2008.0024
- Hur, Y., Ko, Y. J., and Valacich, J. (2007). Motivation and concerns for online sport consumption. *Journal of Sport Management*, 21(4), pp. 521-539. DOI: 10.1123/jsm.21.4.521
- Izquierdo Sanchez, S., Elliott, C., and Simmons, R. (2016). Substitution between leisure activities: a quasi-natural experiment using sports viewing and cinema attendance. *Applied Economics*, 48(40), pp. 3848-3860. DOI: 10.1080/00036846.2016.1145353
- Kagel, J. H., and Roth, A. E. (Eds.). (2016). *The handbook of experimental economics* (Vol. 2). Princeton university press. DOI: 10.1111/1475-4932.12573
- Kahn, L. M. (2000). The sports business as a labor market laboratory. *Journal of Economic Perspectives*, 14(3), pp. 75-94. DOI: 10.1257/jep.14.3.75
- Kotrba, V. (2019). Direct preferences of sports fans: Is there a superstar effect in the fantasy league? *Journal of Behavioral and Experimental Economics*, 78, pp. 89-97. DOI: 10.1016/j.socec.2018.12.010
- Kotrba, V. (2020). Heuristics in fantasy sports: is it profitable to strategize based on favourite of the match?. *Mind & Society*, 19, pp. 195-206. DOI: 10.1007/s11299-020-00231-7
- Kotrba, V. (2021). Nationality-Based Prejudice in Sports: Fan Demand for Foreign Athletes in the Czech Republic. *International Journal of Sport Finance*, 16(1), pp. 3-15. DOI: 10.32731/IJSF/161.022021.01
- Lahvička, J. (2010). Attendance of ice hockey matches in the Czech Extraliga. Available at SSRN 1729845. DOI: 10.2139/ssrn.1729845

- Lee, J., Ruihley, B. J., Brown, N., and Billings A. C. (2013). The effects of fantasy football participation on team identification, team loyalty and NFL fandom. *Journal of Sports Media*, 8(1), pp. 207-227. DOI: 10.1353/jsm.2013.0008
- List, J. A. (2007). Field experiments: a bridge between lab and naturally occurring data. *The BE Journal of Economic Analysis and Policy*, 5(2). DOI: 10.2202/1538-0637.1747
- Lopez-Gonzalez, H., Estévez, A., and Griffiths, M. D. (2017). Marketing and advertising online sports betting: A problem gambling perspective. *Journal of Sport and Social Issues*, 41(3), pp. 256-272. DOI: 10.1177/0193723517705545
- MacDonald, D. N., and Reynolds, M. O. (1994). Are baseball players paid their marginal products?. *Managerial and Decision Economics*, 15(5), pp. 443-457. DOI: 10.1002/mde.4090150507
- Marsden, P. V., and Wright, J. D. (Eds.). (2010). Handbook of survey research. *Emerald Group Publishing*. DOI: 10.3989/ris
- Massari, M. G. (2006). When fantasy meets reality: the clash between on-line fantasy sports providers and intellectual property rights. *Harv. JL and Tech.*, 19(2), pp. 443-465.
- Mobius, M. M., and Rosenblat, T. S. (2006). Why beauty matters. *American Economic Review*, 96(1), pp. 222-235. DOI: 10.1257/000282806776157515
- Moorman, A. M. (2008). Fantasy sports leagues challenged as illegal gambling. *Sport Marketing Quarterly*, 17(4), pp. 232-234.
- Munting, R. (1996). An economic and social history of gambling in Britain and the USA. *Manchester University Press*. DOI: 10.1007/s12117-997-1182-0
- Nesbit, T. M., and King, K. A. (2010a). The impact of fantasy sports on television viewership. *Journal of Media Economics*, 23(1), pp. 24-41. DOI: 10.1080/08997761003590721
- Nesbit, T. M., and King, K. A. (2010b). The impact of fantasy football participation on NFL attendance. *Atlantic Economic Journal*, 38(1), pp. 95-108. DOI: 10.1007/s11293-009-9202-x
- Reiche, D. (2013). The prohibition of online sports betting: a comparative analysis of Germany and the United States. *European Sport Management Quarterly*, 13(3), pp. 293-314. DOI: 10.1080/16184742.2012.741140
- Rosen, S., and Sanderson, A. (2001). Labour markets in professional sports. *The Economic Journal*, 111(469), pp. 47-68. DOI: 10.1111/1468-0297.00598
- Rottenberg, S. (1956). The baseball players' labor market. *Journal of Political Economy*, 64(3), pp. 242-258. DOI: 10.1086/257790
- Schumaker, R. P., Solieman, O. K., and Chen, H. (2010). Predictive modeling for sports and gaming. In *Sports Data Mining* (pp. 55-63). *Springer*. DOI: 10.1007/978-1-4419-6730-5_6
- Schwarz, A. (2004). *The Numbers Game: Baseball's Lifelong Fascination with Statistics*. *Dunne Books*. DOI: 10.1198/000313005X43551
- Scully, G. W. (1973). Economic discrimination in professional sports. *Law and Contemp. Probs.*, 38, pp. 67.
- Smith, V. L. (1976). Experimental economics: Induced value theory. *American Economic Review*, 66(2), pp. 274-279.

- Statista. (2019). North America sports market size from 2009 to 2022. Retrieved from <https://www.statista.com/statistics/214960/revenue-of-the-north-american-sports-market/>
- Sung, H., Mills, B. M., and Mondello, M. (2019). Local Broadcast Viewership in Major League Soccer. *Journal of Sport Management*, 33(2), pp. 106-118. DOI:10.1123/jsm.2018-0022
- Szymanski, S. (2000). A market test for discrimination in the English professional soccer leagues. *Journal of Political Economy*, 108(3), pp. 590-603. DOI: 10.1086/262130
- Tassoni, C. J. (1996). Representativeness in the market for bets on National Football League games. *Journal of Behavioral Decision Making*, 9(2), pp. 115-124. DOI: 10.1002/(SICI)1099-0771(199606)9:2<115::AID-BDM220>3.0.CO;2-B
- Thaler, R., and Ziemba, W. T. (1988). Anomalies: Parimutuel betting markets: Racetracks and lotteries. *Journal of Economic perspectives*, 2(2), pp. 161-174. DOI: 10.1257/jep.2.2.161
- Tversky, A., and Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, 185(4157), pp. 1124-1131. DOI: 10.1126/science.185.4157.1124
- Weaver, R. J. (2010). Online Fantasy Sports Litigation and the Need for a Federal Right of Publicity Statute. *Duke L. and Tech. Rev.*, 9, pp. 1-25.
- Williams, L. V. (1999). Information efficiency in betting markets: A survey. *Bulletin of Economic Research*, 51(1), pp. 1-39. DOI: 10.1111/1467-8586.00069
- Wilson, P. D., and Ying. Y. (2003). Nationality Preferences for Labour in the International Football Industry. *Applied Economics*, 35, pp. 1551-1559. DOI: 10.1080/000368403200010048
- Winters, K. C., Stinchfield, R. D., and Kim, L. G. (1995). Monitoring adolescent gambling in Minnesota. *Journal of Gambling Studies*, 11(2), pp. 165-183. DOI: 10.1007/BF02107113
- Woodland, L. M. and Woodland, B. M. (2015). The National Football League season wins total betting market: The impact of heuristics on behavior. *Southern Economic Journal*, 82(1), pp. 38-54. DOI: 10.4284/0038-4038-2013.145
- Wolfers, J. (2006). Point shaving: Corruption in NCAA basketball. *American Economic Review*, 96(2), pp. 279-283. DOI: 10.1257/000282806777211757