Applications of Sports Data to Study Decision Making*

Loukas Balafoutas a, Subhasish M. Chowdhury b, and Henning Plessner c

a Department of Public Finance, University of Innsbruck, 6020 Innsbruck, Austria
b Department of Economics, University of Bath, Bath BA2 7AY, UK
c Institute of Sports and Sports Sciences, University of Heidelberg, 69120 Heidelberg, Germany

Abstract

The features and determinants of human decision making, especially under competitive environments, have been a central theme of research in economics and psychology alike. This special issue centers on applying sports data to understand behavior and decision making in competitive environments. In particular, we assemble original research with sports data that provides us with insights into hitherto unanswered questions relating to various strategic and psychological aspects of decision making.

Keywords: Sports data; Decision making; Economic psychology
JEL Codes: D9; L8; Z2

* Corresponding author: Loukas Balafoutas (Loukas.Balafoutas@uibk.ac.at).
We thank Carlos Alós-Ferrer, Martin Kocher, and Stefan Schulz-Hardt (the editors), and all the anonymous referees for their helpful comments in preparation of this special issue.
1. Introduction

This special issue of the *Journal of Economic Psychology* brings together sixteen research articles that use data from various sports to study a diverse set of topics in economics, psychology, and management. The common thread of these articles is human behavior and economic decision making by individuals and teams, and the specific topics covered in this special issue include, among others: strategic behavior by teams and individuals in competitive situations, effort provision in teams, task scheduling, psychological momentum effects, judgement and decision biases, life satisfaction, and (strategic) display of emotions. In section 3 of this editorial, we summarize each study and its contribution to our understanding of decision making.

The use of data from sports tournaments in economic and psychological research is not a new phenomenon. In an influential study, Szymbanski (2003) provides a comprehensive review of literature using sports data in order to inform contest theory and design. There are existing books on this topic (e.g., Andreff and Szymbanski, 2006; Leeds et al., 2018), a 2010 special issue published in the *Journal of Economic Psychology* on the economics and psychology of football (see Kocher and Sutter, 2010 for details), and even journals dedicated to the economics or psychology of sports (e.g., *Journal of Sports Economics, Journal of Sport and Exercise Psychology, Psychology of Sport and Exercise*). An overview of psychological approaches to judgment and decision making in sports can be found in Bar-Eli et al. (2011), while the relevance of the findings in this field for economic behavior in general and contests in particular was recently outlined by Ke et al. (2017) and Bar-Eli (2018). However, whereas researchers continue to work on the application of sports data to understand economic decision making, outlets for such research are still relatively sparse.

Over the last decade, this line of research has gained in popularity, as researchers have shown how one can use sports data not only to test economic theories, but also to answer very diverse questions and bring in new insights into human behavior. During this period, a number of researchers have studied various sports, e.g., tennis (Malueg and Yates, 2010; Cohen-Zada et al., 2017; Ely et al., 2017), golf (Brown, 2012), judo (Balafoutas et al., 2012), horse racing (Brown and Chowdhury, 2014), football (Bar-Eli et al., 2007; Palacios-Huerta, 2014; Schreyer et al., 2016; Arrondel et al., 2019), gymnastics (Damisch et al., 2006), and even table football (Memmert et al., 2015), and have shown how sports data can be applied to questions in economics and psychology and how they can open up new questions in behavioral science.
In this special issue we focus specifically on the applications of sports data to study human behavior and decision making. The research included here follows a common thread, but at the same time it is diverse in the ways of examining human behavior. We are confident that this special issue extends our understanding of economic decision making and psychological processes, and complements the existing literature while also providing avenues for future research.

2. Methodological advantages of using sports data

Research using data from sports tournaments has benefited from several advantages, as sports data combine attractive features of laboratory experiments and field data. They ensure a sufficient degree of ‘experimental’ control, given that sports follow clearly documented and strictly enforced rules, while at the same time retaining the external validity of field data. We would summarize these advantages as follows.

First, in sports tournaments there is observability: contrary to many situations in business, politics, or other domains of public and economic life in which individuals can conceal their behavior (or are legally obliged to do so), behavior in sports is observable and very well documented. Most sports events are nowadays televised and recorded, and either the full events or at least informative summaries thereof are widely available to the public. This feature is of particular importance for the analysis of unethical behavior. While field data often suffer from issues such as endogeneity, omitted variable bias, or unobserved heterogeneity, such issues are less prominent with sports data thanks to observability. In addition, observability means that researchers who use sports data often have access to very large datasets, ensuring sufficient power in the statistical analysis.

Second, sports tournaments typically have well-defined rules, which clearly define winners and losers, prizes, illegitimate actions and the consequences of such actions. These rules are set by international regulating bodies and are known to all participants, reducing the likelihood that certain patterns of behavior are driven by confusion. An additional – and very important – ‘bonus’ of this feature is that these rules are occasionally subject to changes (e.g., the introduction and consequent removal of the ‘golden goal’ rule in football), thereby creating natural experiments that can be exploited to estimate causal effects of various environmental aspects (such as incentives). Even when no such exogenous rule changes are introduced, there is abundant naturally occurring randomness in sports tournaments: to give one simple example,
think of coin-tossing to decide which team shoots first in penalty shootouts in football, which team may choose to bat (or field) first in cricket, or which player can choose to serve first or second in a tennis match. In other words, many variables of theoretical interest that need to be manipulated artificially in laboratory studies vary naturally in the world of sports.

These two features, observability and the presence of well-defined, transparent rules, broadly highlight the advantages of sports data over non-experimental field data. However, both features are also present in well-designed experimental studies, especially in laboratory settings that include tight experimental control and full observability. Next, we discuss two features that we consider advantages of sports data over lab experiments.

The third advantage of sports tournaments is that they feature high incentives: if one critique to lab experiments is that stake sizes are typically small, then sports tournaments offer the possibility to examine the behavior of professional athletes and teams under very strong monetary incentives. For example, the singles winners of the 2018 US Open Tennis tournament won $3.8 Million each (even the first round winners secured $93,000) and the winner of the 2018 PGA Championship earned $1.89 million; in team sports, the winning team in the 2018 Indian Premier League cricket earned $2.8 million, and the winning team of the Stanley Cup won $2.66 million in 2018. This money comes on top of the related wage payment, as well as the endorsement and advertising revenues for the players.

A fourth, related advantage of data from sports tournaments is that, unlike laboratory data with student samples in relatively artificial environments, they refer to the behavior of experienced professionals as it is documented in their actual work environment. This enhances the external validity of this kind of data, in the sense that we are dealing with real decisions by professionals, which have large financial consequences.

Fifth, we wish to highlight a general advantage of sports data as a methodological choice, which we call versatility. Versatility is a very prominent feature in this special issue and it refers to the fact that sports data allow us to study a very large number of relevant and important topics in economics, management, and psychology. A non-exhaustive list of such topics in previous literature includes tests of tournament or game theory, strategic effort provision by individuals and teams, studies on unethical behavior (e.g., sabotage, cheating or corruption), risk attitudes, home advantage, and psychological aspects of decision making such as momentum effects, performance under pressure, or bias by game officials.
Versatility also refers to the setting under which the data are collected. As already mentioned above, sports competitions offer rich possibilities to study behavior by individuals or teams. Analyzing data from teams adds one additional layer of complexity, since conflicts of interest or coordination issues among athletes are likely to affect observed behavior. On the other hand, teams are better suited to study topics such as the emergence and impact of leadership, or of managerial decisions. Moreover, sports tournaments feature different formats (such as elimination versus round-robin tournaments, or varying numbers of tournament participants), introducing additional aspects that can potentially be of interest to researchers.

### 3. Overview of the articles in this special issue

The special issue features a total of sixteen articles that use data from ten different individual and team sports: baseball, basketball, football, American football, ice hockey, squash (team sports); chess, cycling, distance running, surfing (individual sports). The data come not only from behavior and observed outcomes during sports tournaments, but also from other domains of the broader sports industry. For instance, two papers consider how salary offers affect player utilization by teams (in football and in baseball), while another informs the optimal design of draft lotteries by showing how the draft regulations may affect club strategies in professional ice hockey.

The analytical methods employed by the authors are also diverse, but a common ground is that they often utilize quasi-experimental random variation in order to make causal inferences. This randomness is either brought about by exogenous changes in the rules of the game or in the institutional environment, or it is the result of naturally occurring variation during a game. In addition, a few of the papers use instrumental variables to estimate causal effects.

#### 3.1. Strategic behavior by teams and individuals in tournaments

Chapsal and Vilain (this issue) test the theory of multiple pairwise battles, where players from two rival teams compete sequentially. Theory predicts that outcomes of past battles should not affect the behavior in the current battle. The authors use data from squash team contests and show that, contrary to the theoretical prediction, winning the first battle indeed increases the probability of winning the subsequent one. They explain this phenomenon with individual utility from winning and contributing to team success.
Brouwer and Potters (this issue) study breakaways, a small set of cyclists who get ahead as a group in cycling races, to investigate the social dilemma within such occasions. This dilemma arises because the opportunity to shelter behind others in a breakaway provides incentives to free-ride. They find an initial positive and concave group-size effect on breakaway success, but the effect eventually turns negative when the number of cyclists in the breakaway becomes ‘too’ large.

An interesting take on incentives to lose in so called ‘dual tournaments’ is found in the article by Fornwagner (this issue). Using data from the National Hockey League, she shows that hockey teams win relatively more games when winning can still increase their chances of reaching the playoff stage, and lose more games, ceteris paribus, when the team has already been eliminated. She further shows that this observation cannot be attributed to lack of motivation, but is due to strategic considerations.

### 3.2. Determinants of individual performance

Pina (this issue) presents evidence from professional surfing tournaments to examine how task scheduling – and in particular, the time elapsed between two tasks – affects performance. The paper exploits the randomness in the scheduling of surfer matchups (due to weather and wave conditions) and finds that a longer time between tasks increases performance and the probability of success in the competition. This result has implications for managerial decisions with respect to assigning and scheduling tasks to workers.

Chan et al. (this issue) examine behavioral changes of players in response to a historical rule change in basketball: the three-point line was shortened for a period of three NBA seasons in the 90ies. Among others, the authors find that players immediately adjusted their 3-point ratio to these changes, although older or more experienced players did not readjust their behavior after the distance was returned to the original condition.

We conclude this sub-section with the ‘ponds dilemma’ (meaning that one has to choose whether to be a big fish in a small pond or a small fish in a big pond), which has been a popular topic of investigation. Zak et al. (this issue) revisit this issue. They use data from chess tournaments in Israel, in which players of medium ranking can choose to compete against stronger opponents in the main tournament or to compete against weaker opponents in the secondary tournament. They find that players choosing to participate in the secondary tournament perform better than the ones choosing to participate in the main tournament.
3.3. Hot hand and psychological momentum

The ‘hot hand’ phenomenon is one of the oldest and most discussed topics in the field. Cotton et al. (this issue) introduce a new analytic tool for studying this effect that is supposed to overcome the problem of conventional methods to overstate the extent of hot hand. By applying this method to a dataset of youth golfers, the authors are able to identify occasional small hot hand effects that may explain the strong belief in the hot hand even when the overall evidence suggests otherwise.

Hot hand can be considered as a special case of the broader concept of psychological momentum, i.e., the experience of a psychological force that enables extraordinary performance within a certain period of time. Two papers in this special issue focus on momentum effects in sports competitions. Dong and Huang (this issue) examine individual performances in a team squash competition and deal with the question of how individuals respond to their own past performance and to that of their teammates. By doing so, the authors can distinguish between strategic and psychological momentum effects, finding more evidence for the former than the latter. Morgulev et al. (this issue) study psychological momentum effects regarding specific overtime situations in basketball. It is often believed that teams that tied the game coming from behind have a higher chance to win in overtime because the momentum is on their side. However, the data from eleven NBA seasons do not reveal such an effect. The psychological momentum appears to be an illusion in this case. In that sense, the findings of Morgulev et al. and Dong and Huang are pointing towards the same direction.

3.4. Judgement and decision biases

The paper by Krawczyk and Wilamowski (this issue) is concerned with the ‘hard-easy’ effect, i.e., the tendency to overestimate (underestimate) the probability of one's success at a task perceived as hard (easy), which relates to the literature on overconfidence (Lichtenstein et al., 1982). They find corresponding evidence for this effect by comparing forecasted finishing times with the race results of more than a million amateur distance runners. Accordingly, the authors suggest that the hard-easy effect is more than a statistical artifact in the world of sports.

Kausel et al. (this issue) apply the idea of outcome bias, i.e., the degree to which the results of a decision affect the judgment of the decision quality, to the evaluation of football players’ performance in games that were decided by a penalty shootout. They find that
reporters’ performance ratings given to about a thousand players were associated with the penalty shootout outcome although it was unrelated to in-game performance. For the first time, this provides evidence for the outcome bias outside of laboratory studies. A similar study of such outcome bias is by Bucciol et al. (this issue) who use the Italian football data to test whether the experience of a win or a loss has an effect of football managers’ risk taking strategy. They find that a single or heavy defeat makes managers more risk seeking, but multiple defeats make them less risk seeking. Interestingly, they find that managers of top teams do not show such bias.

Two papers study the effects of the sunk cost fallacy (Arkes and Blumer, 1985) in sports, presenting largely contradicting evidences. Keefer (this issue) uses Major League Baseball final-offer arbitration to analyze the effects of sunk costs. He finds that a player’s compensation has a direct effect on his plate appearances, even while controlling for beliefs. However, this does not significantly affect individual or team performance. Hackinger (this issue), on the other hand, investigates the effects of sunk costs using data from football and the transfer fees for football players. He finds no effects of the transfer fee on a player’s match appearance at a seasonal level – rejecting any presence of a sunk cost fallacy among football managers. There is, however, a weak effect of such fallacy at the match level that decreases with a player’s tenure.

3.5. Emotions and well-being

Hopfensitz and Mantilla (this issue) use pictures of football players in Panini sticker albums over twelve editions of the World Cup as their dataset, rating each picture with a face-reading software on the extent to which the player displays anger and happiness. The paper documents a positive correlation between both of these emotions and a team’s success in the competition and discusses the extent to which emotion display through facial expressions can be used as a signal in sports competitions.

By crossing data from a nationally representative survey with outcomes in US college American football games, Janhuba (this issue) is able to examine the impact of unexpected results on subjective well-being with a remarkably large sample. In particular, he finds that unexpected wins lead to higher subjective well-being, while unexpected losses do not lead to lower ratings, thus contradicting prior research that has reported effects in both directions.
4. Concluding remarks

In the introduction of this editorial note, we have referred extensively to a number of attractive features of using sports data for the study of decision making (observability, well-defined rules, high incentives, professional participants, versatility) and explained how these features combine methodological advantages of experimental and field data. A brief discussion of limitations is also due. Like with any kind of evidence that is meant to inform policy, concerns about external validity arise when the evidence comes from outside the specific domain to which the policy applies. For instance, if we use data from the behavior of football players within a team and generalize to employee behavior (or misbehavior) at their workplace, the differences between the two settings – with respect to incentives and reward structures, assessment criteria, interaction horizon, effort observability, and many other aspects – must be kept in mind. In addition, findings are often driven by the specifics of a certain sport (e.g., its rules and culture) and cannot automatically be transferred to other sports. It should be noted, however, that this limitation is much less binding whenever sports data are used to test theory, or to identify strategic or psychological effects that are likely to cover behavior across a wide range of circumstances and to display sufficient consistency. A further caveat that must be kept in mind is that, compared to data generated in the laboratory, sports data typically display less experimental control given that variation in the data is either non-random, or the result of exogenous changes outside of the researcher’s control. Hence, while data from sports tournaments combine the advantages of lab and field evidence, they also suffer from certain methodological challenges associated with both, lab experiments (concerns about external validity) and field data (concerns about sufficient control).

The studies in this special issue provide answers to several questions, but also raise a few new ones. They show the need for more sports-specific research regarding behavioral biases (such as the sunk cost fallacy), the need for revisiting theory after the stress test with sports data, and a general need for robustness checks when formulating policy implications. The underlying idea behind this special issue, i.e., using sports data to investigate topics in economics and psychology, remains very relevant for both disciplines and has the potential to generate research of high impact and visibility. We hope this special issue will extend our knowledge to answer existing unanswered questions and intrigue new ones in the area of economic psychology. We also hope that it will further motivate behavioral researchers to exploit data from sports that are not investigated enough till date.
References


