Psychological Pressure in Competitive Environments: Evidence from a Randomized Natural Experiment*

Jose Apesteguia[†] Ignacio Palacios-Huerta[‡]

October 2008

Abstract

Much like cognitive abilities, emotional skills can have major effects on performance and economic outcomes. This paper studies the behavior of professional subjects involved in a dynamic competition in their own natural environment. The setting is a penalty shoot-out in soccer where two teams compete in a tournament framework taking turns in a sequence of five penalty kicks each. As the kicking order is determined by the random outcome of a coin flip, the treatment and control groups are determined via explicit randomization. Therefore, absent any psychological effects, both teams should have the same probability of winning regardless of the kicking order. Yet, we find a systematic first-kicker advantage. Using data on 2,731 penalty kicks from 262 shoot-outs for a three decade period, we find that teams kicking first win the penalty shoot-out 60.5% of the time. A dynamic panel data analysis shows that the psychological mechanism underlying this result arises from the asymmetry in the partial score. As most kicks are scored, kicking first typically means having the opportunity to lead in the partial score, whereas kicking second typically means lagging in the score and having the opportunity to, at most, get even. Having a worse prospect than the opponent hinders subjects' performance. Further, we also find that professionals are self-aware of their own psychological effects. When a recent change in regulations gives winners of the coin toss the chance to *choose* the kicking order, they rationally react to it by systematically choosing to kick first. A survey of professional players reveals that when asked to explain why they prefer to kick first, they precisely identify the psychological mechanism for which we find empirical support in the data: they want "to lead in the score in order to put pressure on the opponent."

*We thank Miguel A. Ballester, Jordi Blanes-i-Vidal, Caterina Calsamiglia, Yeon-Koo Che, Pedro Dal Bo, Manuel E. Mejuto, Patxo Palacios, Steve Pischke, Antonio Rangel, Yona Rubinstein, Giulio Seccia and seminar participants at UCL, Fundación Ramón Areces, IMEBE 2007, Universidad Carlos III, and the IAS in Vienna for helpful comments. Ozan Eksi provided excellent research assistance. We gratefully acknowledge financial support from the Fundación Ramón Areces.

†Universitat Pompeu Fabra. Email: jose.apesteguia@upf.edu.

‡London School of Economics. Email: i.palacios-huerta@lse.ac.uk.

1 Introduction

At least since Hume (1739) and Smith (1759), psychological elements have been argued to be as much a part of human nature, and possibly as important for understanding human behavior, as the strict rationality considerations included in economic models that adhere to the rational man paradigm. Clearly then, any study of human behavior that omits these elements can yield results of unknown reliability.

Much as the rationality principle has successfully accommodated social attitudes, altruism, values and other elements (see, e.g., Becker (1976, 1996), Becker and Murphy (2000)), behavioral economics attempts to parsimoniously incorporate psychological motives not traditionally included in economic models. Theoretical models in this area firmly rely for empirical support on the observation of human decision making in laboratory environments. Laboratory experiments have the important advantage of providing a great deal of control over relevant margins. In these settings, observed behavior often deviates from the predictions of standard economic models. In fact, at least since the 1970s, a great deal of experimental evidence has been accumulated demonstrating circumstances under which strict rationality considerations break down and other patterns of behavior, including psychological considerations, emerge. Thus, an important issue is how applicable are the insights gained in laboratory settings for understanding behavior in natural environments. This challenge, often referred to as the problem of "generalizability" or "external validity," has taken a central role in recent research in the area.[1]

The best and perhaps only way to address this concern is by studying human behavior in real life settings. Unfortunately, however, Nature does not always create the circumstances that allow a clear view of the psychological principles at work. Furthermore, naturally occurring phenomena are typically too complex to be empirically tractable in a way that we can discern psychological elements from within the characteristically complex behavior exhibited by humans.[2]

[1] One concern arises from the fact that "the very control that defines the experiment may be putting the subject on an artificial margin. Even if behavior on that margin is not different than it would otherwise be without the control, there is the possibility that constraints on one margin may induce effects on behavior on unconstrained margins" Harrison and List (2004). A related concern, as expressed for instance in Aumann (1990, 2005), is that in experiments "the monetary payoff is usually very small. More importantly, the decisions that people face are not ones that they usually take, with which they are familiar ... The whole setup is artificial. It is not a decision that really affects them and to which they are used."

[2] See Della Vigna (2007) for a survey of existing work.

In this paper we take advantage of an unusual opportunity. We study a randomized natural experiment, that is a real life situation in which the treatment and control groups are determined via explicit randomization. As is well known, this situation represents a critical advantage in that it guarantees internal validity; that is, it satisfies the conditions for causal inference (Manski, 1995). The subjects in the experiment we study are professionals who have to perform a simple task in a dynamic tournament competition. In soccer, one of the methods of determining the winning team where competition rules require that one team is declared the winner after a drawn match, is by the two teams taking kicks from the penalty mark. This method is used worldwide in all the major elimination tournaments involving both national teams (e.g., World Cups, European Cups, American Cups) and club teams (e.g., Champions League, UEFA Cup). From the time it was first introduced in 1970 until 2003, the basic procedure was as follows:

- Both teams take five penalty kicks;
- The kicks are taken *alternately* by the teams;
- The referee *tosses a coin* and the team whose captain *wins* the toss takes the *first* kick;
- If, after both teams have taken five kicks, both teams have scored the same number of goals, kicks continue to be taken in the same order until one team has scored a goal more than the other from the same number of kicks.

This randomized experiment gives us the chance to study a situation that is familiar to the subjects and in the natural setting where they operate. The subjects are professionals; in fact, among the highest paid professionals in the world, and the task they have to perform (kick a ball once) is one of the simplest they could possibly be asked to perform. Further, as will be discussed in the next section, the setting concerns an important framework of analysis (the tournament model) in labor economics and the economics of organizations. Moreover, from an empirical perspective all the relevant variables are perfectly observable, the task is effortless, outcomes are decided immediately, and with only two possible outcomes (score, no score) risk plays no role in the analysis. Finally, individuals are subject to high incentives, and are therefore interested in performing the best they can. In fact, their actions often have huge consequences not only for their individual careers, but also for their team, their city and even their country as in a World Cup final, for instance.

The explicit randomization mechanism used to determine which team goes first in the sequence, in a situation where both teams have exactly the same opportunities to perform a task, suggests that we should expect the first and second teams to have exactly the same probability of winning the tournament. Yet, we find strongly significant and quantitatively important differences. Using data on 1,343 penalty kicks from 129 penalty shoot-outs over the period 1976-2003, we find that teams that take the first kick in the sequence win the penalty shoot-out 60.5 percent of the time. As these differences in performance arise from the randomly determined differences in the kicking order, the characteristics of the setting are such that they allow us to attribute this average treatment effect to psychological effects resulting from the consequences of the kicking order.

The paper is structured as follows. Section 2 briefly reviews related literature from labor economics on tournaments, the role of emotional skills as determinants of performance and other outcomes, performance under pressure, reference-dependent preferences, and the recent literature modeling confidence and pessimism.

Section 3 describes in detail the setting and the natural experiment, and Section 4 goes over the data and provides the main empirical results of the analysis.

In Section 5 we then try to understand the mechanism whereby teams kicking first are more likely to win. We begin by providing descriptive evidence of the dynamic performance of the subjects. In a first subsection we estimate scoring probabilities using a random effects dynamic panel data model with lagged endogenous variables which accounts for state dependence and unobserved heterogeneity. The results show that lagging in the score is, in fact, what hinders the performance of the subjects. Since most kicks result in goals, kicking first typically means having the opportunity to break the tie and take the lead in the score, whereas kicking second typically means lagging in the score and having the opportunity to, at most, get even. These differences in the state of the competition and prospects at the time the subjects perform their task generate the treatment effect we observe in the data. In a second subsection we provide a discussion of theoretical models that capture what we observe in the data and additional evidence showing that lagging in the score hinders the performance of kickers.

In section 6 we study the change introduced in 2003 in the randomization procedure whereby the winner of the coin toss was no longer required to go first in the sequence but was instead required to *choose* whether to kick first or second. This change in the procedure is important in that it allows us to study (i) whether subjects are aware of the advantage of going first, (ii) whether they rationally respond to it by systematically choosing to kick first, and (iii) whether, when surveyed, they can identify the psychological mechanism for which we find support in the data and attribute to it the reason for their choice. We find that the answers to all three of these questions are affirmative. Consistent with these answers, the patterns that are found in the data for the period 2003-2008 are the same as for the period 1970-2003. Finally, Section 7 concludes.

2 Related Literature

This paper relates to several strands of literature in economics and psychology.

First, the natural setting is that of a tournament. Tournaments are pervasive in organizations. They were formally introduced by Lazear and Rosen (1981), and over the last couple of decades a large literature has studied both theoretically and empirically a number of important aspects of this incentive scheme.[3] Despite the large body of work, however, we are aware of no evidence documenting how psychological or emotional effects may be relevant in explaining the performance of subjects competing in a tournament setting. One possible reason for this is that the difficulty in clearly observing actions, outcomes, choices of risky strategies, and other relevant variables in naturally occurring settings is already exceedingly high, and as a result it is not possible to discern with sufficient precision whether there are, in addition to these variables, any psychological elements at work.

The characteristics of the setting we study, however, are ideal for overcoming these obstacles. Variables such as the choice of effort levels and risky strategies that are typically hard to measure in tournaments and other competitive situations play no role in our setting: the task (kicking a ball once) involves no physical effort and, with only two possible outcomes (score or no score), risk plays no role either.[4] Outcomes

[3]See also Green and Stokey (1983), Nalebuff and Stiglitz (1983), Rosen (1986) for early contributions, and Prendergast (1999) for a review. For empirical work on tournaments see Ehrenberg and Bognanno (1990) in a sports context, and for experimental work Bull, Schotter and Weigelt (1987).

[4]The role of risk in tournament competitions has been studied in Bronars (1987), Hvide (2002) and Hvide and Krnstiansen (2003). In dynamic competition games, there is a literature on the "increasing dominance" effect of a leader over a rival (e.g., Gilbert and Newbery (1982), Cabral and Riordan (1994) and Cabral (2002, 2003)), which studies the strategic amount of resources to use and their allocation (i.e., the strategic choice of variance and covariance) throughout a competition.

can be perfectly observed and are immediately determined after players make their choices; that is, there is no subsequent play. The fact that there is no subsequent play is important, indeed critical, to establish the empirical results.[5] Further, the rules of the competition are precisely established and require that subjects be always in the same physical situation (same position and location). Viewed from this perspective, the setting offers substantive advantages to study the role of psychological elements in competitive environments.

Second, to the extent that the psychological or emotional effects we study are endogenous to the state of the competition itself, the characteristics of the setting are valuable for understanding the determinants of performance not only in tournaments but, more generally, in competitive settings. Heckman (2008) offers a thorough survey indicating that emotional skills can be important determinants of socioeconomic outcomes, contribute to performance at large, and even help to determine cognitive achievement. They are, however, hard to document in natural settings.

Third, there are some recent models of preferences which bear on the analysis. Köszegi and Rabin (2006) develop a model of reference-dependent preferences where a person's reference point is her rational expectations about outcomes, and "gain-loss" utility evaluations around this point influences her behavior. In our setting, the score at the time a player has to perform his task (the state of the competition) appears to act as a reference point. It is then the gain-loss (or "ahead-behind") asymmetry associated with this partial score that has an impact on behavior. Put differently, the psychological mechanism is one in which differences in reference points and "local" prospects appear to have a differential impact on behavior.

Relatedly, Epstein and Kopylov (2007) develop an axiomatic model of pessimism where individuals lose confidence in their outlook as they approach the moment of truth. Essentially, in their model, the "pessimistic belief" varies with the prospect in hand, and this is achieved endogenously. In the context of our setting, the moment of truth is different for different partial scores in that the prospect of getting ahead in the score is better than the prospect of merely getting even. Beliefs then map into

[5]The reason is that if there were subsequent actions that contribute to determine the outcome, we would need to have detailed information on the subjects' choice of effort levels and choice of risky strategies in those actions. Further, the extent to which subjects' may have asymmetric information concerning their effort levels and heterogeneity in risk attitudes may also act as relevant determinants. These aspects mean that situations in which a coin toss is used to decide, for instance, initial sides or which team begins play (see, e.g., Bhaskar (2008)) are an order of magnitude more complex and untractable than a penalty kick in order to study the presence of psychological elements.

suboptimal actions or "trembling feet." An interesting aspect of Epstein and Kopylov's analysis is that the individual is sophisticated and forward-looking, in that he is fully aware that he may develop cold feet as the moment of truth approaches. They indicate, however, that they are "not familiar with definite evidence on whether individuals are self-aware to this degree." Interestingly, as mentioned earlier, this open question can be addressed using the opportunity provided by the change introduced in 2003 in the procedure used to determine the order in the sequence. Since then, winners of the coin toss must choose the kicking order. If they were sophisticated and forward-looking, the differences in their degree of cold feet imply that they should systematically choose to go first. Our data confirms that this is in fact the case.

Fourth, an important literature in social psychology has studied expert performance and performance under pressure such as that induced by high stakes, the presence of an audience and others (see, for instance, Ericsson et al (2006), Beilock (2007) and other references therein). Ariely, Gneezy, Loewenstein and Mazar (2008) review and discuss this literature in the context of an study of whether increases in motivation and effort result in improved performance. In our setting, however, *both* teams have the *same* stakes and both perform in front of the *same* audience, an audience which in many shoot-outs supports roughly equally both teams. More importantly, although different forms of pressure may be complements with each other, the explicit randomization procedure that is used means that there is no reason why one team should be systematically more affected by the stakes or the audience than the other. The novel result we obtain from the perspective of this literature is that differences in the interim state of the competition caused by the kicking order generate differences in the psychological pressure that drives the effects on performance that we observe. We will not especulate as to the actual form that these psychological differences may take beyond indicating that they may be associated with mechanisms such as increased arousal, greater shifting of mental process from "automatic" to "controlled," or differences in the narrowing of attention (see Ariely, Gneezy, Loewenstein and Mazar (2008), Kahneman (1973)).

In contrast to the size of the psychology literature, the economics literature on psychological effects on economic decision making is fairly limited, with pioneering theoretical contributions by Loewenstein (1987), Caplin and Leahy (2001) and Rauh and Seccia (2006) on anxiety and anticipatory emotions. We are aware, however, of no empirical contributions with evidence from strictly competitive environments in real life. In terms of this literature, our results may be attributed to differences in cognitive anxiety, a term that is defined as a mental component involving "negative expectations and cognitive concerns about oneself, the situation at hand, and potential consequences" (Morris et al., 1981, p. 541).

Lastly, there is some economic literature on the ex post fairness of certain regulations in sports where a coin flip that determines the order of play may have a significant impact on the outcome of a game by giving the winner of the coin flip *more* chances to perform a task (see, for example, Che and Hendershott (2007) and Wall Street Journal (2003) for the case of extra-time sudden-death regulations in the National Football League). Our results show that, even under ideal circumstances where a coin flip determines only the order of competition and both teams have exactly the *same* chances, human nature is such that the outcome of a perfect randomized trial may be considered ex post unfair.[6]

3 The Randomized Natural Experiment

A *penalty shoot-out* is simply a sequence of *penalty kicks*. The rules that govern a penalty kick are described in the *Official Laws of the Game* (FIFA, 2007) of the world governing body of soccer, the *Fédération Internationale de Football Association*.[7] In a penalty kick, the positions of the ball and the players are determined (FIFA, 2007, p. 45) as follows:

- "The ball is placed on the penalty mark in the penalty area.
- The player taking the penalty kick is properly identified.
- The defending goalkeeper remains on the goal line, facing the kicker, between the goalposts, until the ball has been kicked.
- The player taking the penalty kicks the ball forward.
- A goal may be scored directly from a penalty kick."

Each penalty kick involves two players: a kicker and a goalkeeper. In the typical kick the ball takes about 0.3-0.4 seconds to travel the distance between the penalty mark and the goal line, which is less than the reaction time plus goalkeeper's movement time to possible paths of the ball. Hence, both kicker and goalkeeper must

[6] In settings such as ours, an ex-post fair regulation would then require both teams to perform their task simultaneously, rather than sequentially.

[7] See Law 1 in FIFA (2007) for details concerning the field of play, penalty area, goals, distances, etc.

move simultaneously and the outcome is determined immediately. The penalty kick has only two possible outcomes: score or no score. There are no second penalties or any form of subsequent play in the event of a goal not being scored. The task can be considered, by any reasonable metric, effortless and with only two possible outcomes risk plays no role.[8] Further, players' actions and outcomes can be perfectly observed. The initial location of both the ball and the goalkeeper is always the same: the ball is placed on the penalty mark and the goalkeeper positions himself on the goal line.

As indicated above, a penalty shoot-out is a sequence of penalty kicks, the purpose of which is to decide the winning team where competition rules require one team to be declared the winner after a drawn match. The official rules and regulations in a penalty shoot-out are given in Appendix A. The shoot-out was first introduced in 1970, and until July 2003 the main characteristics were as follows:

• "The referee tosses a coin and the team whose captain wins the toss takes the first kick.

• The referee keeps a record of the kicks being taken.

• Subject to the conditions explained below, both teams take five kicks.

• The kicks are taken alternately by the teams.

• If, before both teams have taken five kicks, one has scored more goals than the other could score, even if it were to complete its five kicks, no more kicks are taken.

• If, after both teams have taken five kicks, both have scored the same number of goals, or have not scored any goals, kicks continue to be taken in the same order until one team has scored a goal more than the other from the same number of kicks."

In July 2003, FIFA decided to change slightly the first regulation in the procedure by replacing it with (italics added):

• "The referee tosses a coin and the team whose captain wins the toss *decides* whether to take the first or the second kick."

The clarity of the rules of a penalty shoot-out, as well as the characteristics and the detailed structure of a penalty kick, present notable advantages for conducting empirical research. The focus of our analysis is the period 1970-2003 where we have a perfect, explicit randomized experiment. Further, we will also use data after 2003 to study the decisions that players make and the implications of these.

[8]We refer here to physical effort, which is typically conceived as a choice variable. With regard to mental effort, arousal is the brain's way of increasing its level of effort, and it is not ordinarily under volitional control (see, e.g., Kahneman, 1973).

4 Data and Empirical Evidence

The data come from the *Union of European Football Associations* (UEFA), the Rec.Sport.Soccer Statistics Foundation, the Association of Football Statisticians, and the Spanish newspaper MARCA. The dataset comprises 262 penalty shoot-outs with 2,731 penalty kicks over the period 1970-2008. It is comprehensive in that it includes *all* the penalty shoot-outs in the history of the main international competitions for national teams (e.g., World Cup, European Championship, American Cup) and club team competitions such as Champions League and the UEFA Cup. It also includes data on national club competitions such as the Spanish Cup, German Cup, and the English F.A. Cup. Table 1 provides a summary.

[Table 1 here]

For every shoot-out of every competition we have information on the date, the identity of the teams kicking first and second, the final outcome, the outcomes of each of the kicks in the sequence (with the exception of one shoot-out), and the geographical location of the game (that is, whether the game was played in a home ground, a visiting ground, or in a neutral field). As just indicated above, the focus of our analysis is the period 1970-2003, and the post-2003 data will be used to assess other relevant aspects.

As is well known, and following the description in Manski (1995), let y_z be the outcome that a subject (a team in our case) would realize if he or she were to receive treatment z, where $z = 0, 1$. Let $P(y_z|\, x)$ denote the distribution of outcomes that would be realized if all subjects with covariates x were to receive treatment z. The objective is to compare the distributions $P(y_1|\, x)$ and $P(y_0|\, x)$. When the treatment z received by each subject with covariates x is statistically independent of the subject's outcomes, we have $P(y_z|\, x) = P(y_z|\, x, z = 1) = P(y_z|\, x, z = 0)$ for $z = 0, 1$. Now let $y \equiv y_1 z + y_0(1 - z)$ denote the outcome actually realized by a member of the population, namely, y_1 when $z = 1$ and y_0 when $z = 0$. Note that $P(y|\, x, z = 1) = P(y_1|\, x, z = 1)$ and $P(y|\, x, z = 0) = P(y_0|\, x, z = 0)$. Hence, if we denote by B the specified set of outcome values (that is, simply win or lose in our case), when the treatment is independent of outcomes, the estimate of the treatment effect $T(B|\, x)$ is simply:

$$T(B|\, x) = P(y \in B|\, x, z = 1) - P(y \in B|\, x, z = 0).$$

Next, we first confirm the statistical similarity of the pre-treatment characteristics of the two teams involved in a shoot-out. The main covariates we are interested in are variables that measure the quality of the teams, their previous experience in shoot-outs, and environmental factors such as the nature of the crowd in the stadium since these may represent differences in support or pressure experienced by the teams. With respect to the quality of the teams, FIFA and UEFA publish yearly rankings both for national teams and clubs based on their performance in certain competitions. For the national team competitions we use the "FIFA rankings," and for international club competitions the "UEFA team rankings."[9] For club competitions at the national level we consider the division or category to which the teams belong at the time of the shoot-out and, when they belong to the same division, their standings in the table at the time of the shoot-out. With respect to experience, we compute the number previous shoot-outs observed in our dataset in which a team has been involved. Lastly, we consider whether a team is playing at its own stadium in front of mostly a supporting home crowd, at the stadium of its opponent in front of a predominantly unfriendly crowd, or at a neutral venue. Table 2 reports the differences in these characteristics.

[Table 2 here]

Consistent with the randomization procedure used to determine the order of play, it is not possible to reject the null hypothesis of equality in any of these covariates at the usual significance levels.

We will now turn to the main result of this paper. As indicated earlier, the estimate of the average treatment effect is simply $P(y \in B|\, x, z = 1) - P(y \in B|\, x, z = 0)$. We compute this effect and find that teams kicking first in the sequence win the penalty shoot-out 60.5% of the time. That is, kicking first conveys a strongly significant (at the 1.7 percent level) and sizeable advantage.

[Figure 1 here]

[9]The methodology used to construct these rankings is described in www.fifa.com and www.uefa.com.

In Table 3 we use a regression framework to provide an estimate of the treatment effect using various probit and logit specifications.

[Table 3 here]

We find, not surprisingly, that the order of play is strongly significant in every specification. Further, it is also interesting to note that none of the covariates that we consider are significant in any of the specifications. These results confirm the significant and sizeable advantage gained by the team that is first to kick. In the next section we turn our attention to trying to understand the mechanism that generates this advantage.

5 Understanding the Mechanism

We begin this section by providing descriptive evidence of winning frequencies and scoring probabilities by round and partial score. We then estimate scoring probabilities using a dynamic panel data model of the performance of the two teams throughout the tournament. This model accounts for unobserved heterogeneity and state dependence.

Table 4 reports round-by-round data of winning frequencies for each team. As indicated in Appendix A, the regulations establish that "if, before both teams have taken five kicks, one has scored more goals than the other could score, even if it were to complete its five kicks, no more kicks are taken." They also indicate that when the shoot-out remains tied after 5 rounds "kicks continue to be taken in the same order until one team has scored a goal more than the other from the same number of kicks." This means that the performance of the teams in each and every round alone after the first 5 rounds may be (when one team scores and the other does not) entirely decisive of the final outcome.

[Table 4 here]

Most of the shoot-outs end in 5 rounds or less, and the rest move into decisive rounds. The team kicking first wins 65.9% of all the shoot-outs that end in 5 rounds or less, and 55.5% of the rest. The lower advantage in the decisive rounds may simply reflect that it is only the teams that fail to capitalize on the advantage that kicking first provides during five rounds which must play these rounds.

Figures 2.1 and 2.2 report the unconditional scoring rates per round for the first 5 rounds, and the unconditional frequencies with which a given team (first or second) is ahead of its opponent in the score *at the end* of each of these rounds.

[Figures 2.1 and 2.2 here]

The scoring rate in the aggregate data is 73.1 percent, 76.3 percent for the first team and 69.7 percent for the second. These rates are lower than the average scoring rate in penalty kicks in the normal course of a game (that is, not in shoot-outs), which is about 80 percent, but similar to the scoring rate in games with a close score (a tie or a one goal difference) and there is little time left to play in the game (see, e.g., Palacios-Huerta (2003)). This would appear to reflect the increased pressure associated with the fact that scoring a goal or not will be a critical determinant of the final outcome both in a shoot-out and in close games.

Figure 2.1 shows that in every round the scoring rate is always greater for the first team than for the second team, while the scoring rate for both teams appear to decline in the later rounds. Figure 2.2 shows that these differences make the first team more likely to be leading in the score than the second team at the end of every round. The difference between the teams is 7 percentage points in each of the first two rounds, not significant at conventional levels, and increases in magnitude in the subsequent rounds, to 12, 13 and 19 percentage points respectively, becoming statistically significant (at the 1 percent level). Interestingly, the frequency with which the first team leads in the partial score *relative* to that of the second team is around 60% greater on average, increasing slightly in rounds 4 and 5 relative to the first three rounds. This suggests that the detrimental effects on performance become more pronounced as the final rounds are approached.

Table 5 provides a more detailed description of both scoring probabilities and winning frequencies by team, round and partial score.

[Table 5 here]

Since most penalty kicks are scored, it comes as no surprise to find that most of the observations for the first team are when the partial score is tied, and for the second team when it is lagging in the score. The scoring rate along these two paths of observations (columns 2 and 4 in the table) is nearly always higher for the first team than for the second, and the same is true if we condition on the same partial score. If we compare these two paths, the scoring rate drops quite significantly for the second team but not for the first. For the second team it falls from about 75-80 percent in the first two rounds to about 62-66 percent in rounds 3 to 5, whereas for the first team it remains fairly stable in the range of 72-78 percent.

The percentage of times with which the teams observed at every round-score combination eventually win the shoot-out reveals that the relative impact of scoring versus not scoring increases over the rounds for both teams. In round 1, for instance, the first team begins with a 60.2 percent chance of winning. If it scores, the probability increases 7.1 percentage points (to 67.3 = 100 - 32.7) and if it misses the probability drops 26.9 percentage points (to 33.3 = 100 - 66.7). The corresponding figures for round 5 are +17.6 and -35.7 percentage points, respectively. Thus, the cumulative impact of any scoring rate differentials over five rounds can be substantial. If by round 5 the score remains even for the first team, its probability of winning drops to 52.9 percent.

In the next subsection we study whether the patterns that appear to be present in the raw data are substantiated in a more rigorous analysis of the dynamics of the tournament.

5.1 Dynamic Panel Data Analysis

In order to understand the underlying mechanism we need to estimate scoring probabilities using a dynamic discrete choice panel data model with lagged endogenous variables that controls for state dependence and unobserved heterogeneity. Given that we need to account for whether the outcome of a penalty kick (score, no score) may be affected by the state of the shoot-out and by past outcomes, we need to deal with regressors that are predetermined but not exogenous. Thus, the outcome may depend on certain intrinsic characteristics of the teams and the penalty shoot-out, the specific sequence of past outcomes, and the state of the tournament shoot-out.

A number of difficulties arise when estimating binary choice panel data models with predetermined variables and unobserved heterogeneity. For instance, parameter estimates from short panels jointly estimated with individual fixed effects can be seriously biased and inconsistent when the explanatory variables are only predetermined as opposed to strictly exogenous (see Arellano and Honoré (2001) for a review).[10] In

[10]In linear models with additive effects, the standard response is to consider instrumental-variables

order to control for the effect of state dependence appropriately in our setting, we estimate in this subsection a semi-parametric, dynamic random effects, discrete choice, panel data model based on Arellano and Carrasco (2003). These authors develop a consistent random effects estimator where: (a) explanatory variables are predetermined but not strictly exogenous, and where (b) individual effects are allowed to be correlated with explanatory variables. This estimator contains a non-parametric conditional expectation of the effects given the predetermined variables, but is otherwise parametric. This makes the estimation of the model affordable without restricting the estimates of the effects by imposing an arbitrary distribution of the conditional expectation.

The basic idea of the model is to define conditional probabilities for every possible sequence of realizations of the state variables. In this sense, we can deal with regressors that are predetermined but not exogenous. Then, the estimator computes the probability of a given outcome along every possible path of past realizations of the endogenous regressors. The panel data structure allows us to identify the effect of individual unobserved heterogeneity since outcomes can be different even when teams share the same history of realizations of the state variables.

Consider two discrete outcomes (score, no score) denoted $y_{it} = \{1, 0\}$. The probability of each of them depends on the specific sequence of past outcomes and the state of the shoot-out tournament. Since outcomes can be different, different experiences change the information set and the expected realizations of future outcomes. To be more specific, the probability of a given outcome may depend on certain intrinsic characteristics of the teams involved in the shoot-out, as well as on their expectation on the realization of the final outcome. This can be written as follows:

$$y_{it} = \mathbf{1}\left\{\beta z_{it} + E\left(\eta_i \mid w_i^t\right) + \varepsilon_{it} \geq 0\right\},$$

$$\varepsilon_{it} \mid w_i^t \sim N\left(0, \sigma_t^2\right),$$

where z_{it} includes the set of time–invariant characteristics of the teams and the shoot-out, x_{it}, plus the state of the shoot-out and the previous outcomes $y_{i(t-1)}$.

estimates which exploit the lack of correlation between lagged values of the variables and future errors in first differences. In non-linear models, however, very few results are available. For fixed effects the few available methods are case-specific (logit and Poisson) and, in practice, lead to estimators that do not converge at the usual $\sqrt{n}$-rate. In the case of random effects, the main difficulty is the so-called initial conditions problem: if one begins to observe subjects after the "process" in question is already in progress, we need to isolate the effect of the first lagged dependent variable from the individual-specific effect and the distribution of the explanatory variables prior to the sample.

We denote by $w_i^t = \{w_{i1}, ..., w_{it}\}$ the history represented by a sequence of realizations $w_{it} = \left\{x_{it}, y_{i(t-1)}\right\}$, and by η_i an individual effect (future outcome realization for team i) whose forecast is revised each period t as the information summarized by the history w_i^t accumulates.[11] The conditional distribution of the sequence of expectations $E(\eta_i \mid w_i^t)$ is left unrestricted, and hence the process of updating expectations as information accumulates is not explicitly modeled. This is the only aspect that makes the model semi–parametric. Given the history of past outcomes, since errors are normally distributed, the conditional probability of $y_{it} = 1$ at time t for any given history w_i^t is:

$$\Pr\left(y_{it} = 1 \mid w_i^t\right) = \Phi\left[\frac{\beta z_{it} + E(\eta_i \mid w_i^t)}{\sigma_t}\right].$$

Since the model has discrete support, any individual history can be summarized by a cluster of nodes $j = 1, ..., J$ representing the sequence of realizations for each vector of characteristics. Thus, the conditional probability can be rewritten as:

$$p_{jt} = \Pr\left(y_{it} = 1 \mid w_i^t = \phi_j^t\right) \equiv h_t\left(w_i^t = \phi_j^t\right), \qquad j = 1, ..., J.$$

The estimation relies on an intuitive idea. In order to remove the unobserved individual effect, we account for the proportion of teams with identical characteristics and history up to time t that realize a given outcome at time t. We then repeat this procedure for every cluster of combinations of demographics and histories in our data. For each cluster we compute the percentage of times that outcome $y_{it} = 1$ occurs. This provides a simple estimate of the unrestricted probability $\hat{p}_{jt}$ for each possible history in the sample. Then, by taking first differences of the inverse of the equation above we get:

$$\sigma_t \Phi^{-1}\left[h_t\left(w_i^t\right)\right] - \sigma_{t-1}\Phi^{-1}\left[h_{t-1}\left(w_i^{t-1}\right)\right] - \beta\left(x_{it} - x_{i(t-1)}\right) = \xi_{it},$$

and, by the law of iterated expectations, we have:

$$E\left[\xi_{it} \mid w_i^{t-1}\right] = E\left[E\left(\eta_i \mid w_i^t\right) - E\left(\eta_i \mid w_i^{t-1}\right) \middle| w_i^{t-1}\right] = 0.$$

This conditional moment condition serves as the basis of the GMM estimation of parameters β and σ_t (subject to the normalization restriction that $\sigma_1 = 1$). Arellano

[11] The specification of Arellano and Carrasco (2003) is more general in the sense that it also includes a time-varying component, γ_t, common to all individuals. In our case all "demographic" variables are time–invariant.

and Carrasco (2003) show that there is no efficiency loss in estimating these parameters by a two–step GMM method where in the first step the conditional probabilities p_{jt} are replaced by unrestricted estimates $\hat{p}_{jt}$, which in our case are the proportion of teams with given characteristics and a given history. Then:

$$\hat{h}_t\left(w_i^t\right) = \sum_{j=1}^{J} \mathbf{1}\left\{w_i^t = \phi_j^t\right\} \cdot \hat{p}_{jt},$$

can be used to define the sample orthogonality conditions of Arellano–Carrasco's GMM estimator:[12]

$$\frac{1}{N}\sum_{i=1}^{N} d_{it}\left\{\sigma_t\Phi^{-1}\left[\hat{h}_t\left(w_i^t\right)\right] - \sigma_{t-1}\Phi^{-1}\left[\hat{h}_{t-1}\left(w_i^{t-1}\right)\right] - \beta\left(x_{it} - x_{i(t-1)}\right)\right\} = 0, \quad t = 2, ..., T,$$

where d_{it} is a vector containing the indicators $\mathbf{1}\left\{w_i^t = \phi_j^t\right\}$ for $j = 1, ..., J$.

This model offers valuable advantages over the very few alternative approaches available. Furthermore, our short panel fits the identification requirements of their GMM estimator.[13] We estimate this model and collect the results in Table 6.

[Table 6 here]

We find that the main determinant of the scoring rate is "partial score -1." It has a negative effect that is strongly significant at conventional levels, regardless of whether or not we include other endogenous variables relating to the state of the shoot-out. This means that lagging in the score hinders the performance of the subjects. Consequently, the team more likely to find itself with a partial score of -1 will have significantly greater chances of losing the tournament. We also find that this effect is mitigated if the kicking team is the one kicking in second place. This

[12] We use the orthogonal deviations suggested by Arellano and Bover (1995) instead of first differences among past values of the state variables.

[13] Alternative fixed-effects approaches, such as Honoré and Lewbel (2002) and Honoré and Kyriazidou (2000), are also far more demanding in terms of data. In particular, they require the exogenous regressors to vary over time, something that does not occur in our data. Honoré and Kyriazidou (2000) include one lagged dependent variable but require that the remaining explanatory variables should be strictly exogenous, thus excluding the possibility of a lagged dependent regressor. Furthermore, their estimator does not converge at the usual $\sqrt{n}$–rate. Honoré and Lewbel (2002) allow for additional predetermined variables but at the cost of requiring a continuous, strictly exogenous, explanatory variable that is independent of the individual effects. See Fernandez-Val (2008) for a characterization of the bias of fixed effect estimators in nonlinear panel data models.

is consistent with the intuition that, while a negative partial score is bad news, it is especially bad news for a team that has had exactly the same opportunities to score as its opponent (as is the case for the first team to kick but not for the second).

Other interesting results in the more complete specification of the third column are that the arguably greater nervousness associated with the decisive rounds (those played after the first set of five penalties) appears to have a negative impact on the probability of scoring for both teams, though only at the 20 percent level. Once we control for the effects of the partial score in the data, none of the exogenous variables other than "Final Game" has a significant effect at conventional significance levels.

With respect to the magnitude of the effects, the marginal effects associated with the transition among different states can be computed as follows. Arellano and Carrasco (2003) show that the probability of a given outcome when we compare two states $z_{it} = z^0$ and $z_{it} = z^1$ changes by the proportion:

$$\hat{\triangle}_t = \frac{1}{N}\sum_{i=1}^{N}\left\{\Phi\left(\hat{\sigma}_t^{-1}\hat{\beta}\left(z^1 - z_{it}\right) + \Phi^{-1}\left[\hat{h}_t\left(w_i^t\right)\right]\right) - \Phi\left(\hat{\sigma}_t^{-1}\hat{\beta}\left(z^0 - z_{it}\right) + \Phi^{-1}\left[\hat{h}_t\left(w_i^t\right)\right]\right)\right\}.$$

Since this proportion depends on the history of past ω_i^t, these marginal effects are different for each partial score in the sample, and for each team. Table 7 presents various marginal effects evaluated in each of the different rounds. In Panel A we report the marginal effects associated with different partial scores, and in Panel B the marginal effects associated with different rounds.

[Table 7 here]

The results in Panel A show that the transition of a team from a partial score of 0 to +1 has a positive impact. For the team kicking first, the increase in the probability of scoring is around 1.50% per round for rounds 2 to 4, reaching 3.7% in the 5th round. The impact is, as expected, greater for the second team, and ranges from 2.37% to 3.52% for rounds 2 to 4. Moving a team from a partial score of 0 to -1 has a negative impact, whose magnitude in absolute value is greater for the team kicking first and lower for the team kicking second at any given round than in the case when we move a team from 0 to +1.

It should be noted that, since we are dealing with a two agent zero-sum game, these marginal effects must be compounded in a zero-sum fashion (when one team goes from 0 to +1, the other team *must* go from 0 to -1) in order to gain a sense of their impact on a team's chances of winning the tournament. Consistent with the basic intuition from the raw data, this compounded effect is greater in rounds 4 and 5 than in the earlier rounds. For the decisive rounds, the marginal effect for the second team (the only one that exists in these rounds) is 6.97%. It is sizeable, although somewhat smaller than the compound effect in round 5.

Lastly, in Panel B we report the marginal effects of kicking first rather than second net of other effects. This effect is positive, though small in magnitude, ranging from 0.28% to 1.40% for the first five rounds, and rising to 2.52% in the critical rounds.

5.2 Discussion

We interpret these results as indicating that lagging in the score, that is the state of being in a worse situation and consequently having a inferior outlook, has a detrimental effect on performance. This is the main effect we find in the data. Consequently, the mechanism that translates this effect into the significantly greater probability of the first team winning the tournament documented in the previous section is that, since most penalty kicks are scored, kicking first typically means having the chance to lead in the partial score, whereas kicking second typically means lagging in the score and having the chance to, at most, get even.

We also find that kicking second has a detrimental effect on performance, although it is small in magnitude. It may be conjectured that, ceteris paribus, the situation is more critical in all of the second team's penalty shots (since there are no more penalties in the round) and this may increase the pressure on the player that has to take the kick.

From a theoretical perspective the main results suggest that a tournament model with just two parameters, say p and q (where $p > 0$ is the probability of scoring when the partial score is positive or zero, and $q > 0$ the probability of scoring when the partial score is negative, with $p > q$ to reflect the psychological pressure), might be useful to capture the main features of the data in a parsimonious way. This model could also be generalized in a number of dimensions such as, for instance, to four parameters p_1, p_2, q_1, q_2 to account for differences in the kicking order per se, in addition to the partial score. In Appendix B we study the $p - q$ model and find it to be the simplest model that *always* generates a first mover advantage. We also briefly discuss various extensions which, depending on the configuration of the parameters, may generate either a first team or a second team advantage, as well as the strategic placement of players when they are heterogeneous in quality.

It is important to remark that, thus far, we have attributed the effects on the performance in a penalty kick to the kicker. A penalty, however, involves 2 players: a kicker and a goalkeeper. And hence it is theoretically possible to consider that it is not that the kicker's performance is hindered by a negative partial score, but rather that the goalkeeper's performance is enhanced when his team's partial score is positive. That is, it is possible that goalkeepers may somehow manage to save more penalties when their teams are leading in the score. Although this is theoretically possible, the following three pieces of evidence indicate that it is unlikely:

(i). For a subset of all the penalty shoot-outs in the sample we have detailed information on whether the no-goals are due to "saves" by the goalkeeper or "misses" by the kicker. We have estimated the same Arellano-Carrasco model for a multinomial logit specification with goals, misses and saves. The results are collected in Table 8.

[Table 8 here]

Panel A reports the raw data in scoring, misses and saving rates for the first and second team. The data show that both teams have basically the same proportion of saves, and hence that the difference in scoring rates between the first and the second team basically corresponds to their difference in misses.

In Panel B we report the results of different regression specifications. We find that the coefficient on "Partial score -1" is positive and highly significant (beyond the 1 percent level) for misses, but insignificant for saves in all the specifications. This means that lagging in the score predicts more misses by the kicker but no more saves by the goalkeeper. The interaction with the "Second team" variable is negative and significant, which means that when the partial score is -1 the first kicking team is more likely to miss. This, as in Table 6, is likely the result of being in a worse situation than the second team (it has had the same number of chances of scoring, whereas, at every kick, the second team has always had one less chance). No variable except the constant term is significant for saves.

We take these results as indicating that the decrease in the scoring rate documented earlier for the second team, which is the one more likely to be behind in the partial score, can be attributed to an increase in misses by the kicker rather than to saves by the goalkeeper of the opposing team.

(ii). We have studied penalty kicks in situations characterized by scoring rates similar to those in penalty shoot-outs (Palacios-Huerta, 2003). These are penalties in regular league games when there are fewer than 5 minutes left in the game and the score is either tied or one team is ahead by one goal. Consistent with the above findings, the results (not reported) of both probit and logit specifications indicate that lagging in the score helps predict more misses but does not predict more saves.

(iii). Lastly, the survey of professional players that we will discuss in the next section indicates that professional players themselves systematically report that the effect is to put pressure on the kicker. Not a single player mentions the possibility that the performance of the first team's goalkeeper may be enhanced when the partial score is in his favor. This is also consistent with the fact that goalkeepers have a small impact, relative to that of kickers, on the outcome of a penalty kick.

6 Are Professionals Aware of the Psychological Effects on Performance?

In July 2003, FIFA introduced a slight change in the procedure used to determine the kicking order. The part of the procedure that establishes that:

- "The referee tosses a coin and the team whose captain wins the toss takes the first kick,"

was replaced by (italics added):

- "The referee tosses a coin and the team whose captain wins the toss *chooses* whether to take the first kick or the second kick"

This change allows us to study the response of professionals to the psychological phenomenon we have documented: (1) Professionals may or may not be aware of it; (2) If they are, they may or may not react optimally to it; (3) And if they do, they may or may not do it for the right reason.

Clearly, if subjects are aware of the psychological effect caused by the order of play and its detrimental impact on performance, they should always choose to go first in the tournament. Unfortunately, there are no public records of players' choices because FIFA regulations do not require referees to record this information (see Appendix A).

In order to get to know what their choices are, we have done the following:

1. First, we watched several videos of matches that ended in a penalty shoot-out. Although the interval between the end of a game and the beginning of a shoot-out is typically used by TV channels to air commercials, it is sometimes possible to catch the very instant when the referee flips the coin and talks to the winner of the toss. For instance, this was the case in the 2006 World Cup final between Italy and France and in the 2008 Champions League final between Chelsea and Manchester United. On both these occasions, the winners of the coin toss (Fabio Cannavaro for Italy and Rio Ferdinand for Manchester United) chose to have their teams kick first. Consistent with their behavior, we have observed that in each and every case, with one exception, the winner of the coin toss chooses to kick first.[14]

2. Second, as an indirect test, if in fact the choice is always or nearly always to kick first, we should see basically no differences between the 1970-2003 and the 2003-2008 data. Consistent with this hypothesis, we studied the 2003-2008 data separately and obtained no significant differences in any of the results with respect to those reported for the 1970-2003 period. Our main finding, the significant advantage gained by the team that kicks first, is 58.6%-41.4% in the 2003-2008 data and 59.5%-40.5% for the entire 1970-2008 period.

3. Our third strategy is perhaps even more conclusive. We conducted a survey of more than 240 players and coaches in the Spanish leagues, both professional and amateur, who were asked the following question: `‘‘Assume you are playing a penalty shoot-out. You win the coin toss and have to choose whether to kick first or second. What would you choose: first; second; either one, I am indifferent; or, it depends?’’` The results are collected in Table 9.

[Table 9 here]

We found that just about 100% of the subjects answered that they would prefer to go first. More importantly, when asked them to explain their decision, they systematically argued that their choice was motivated by the desire to put pressure on the kicker of the opposing team. Coding their answers to a second question we asked: `‘‘Please explain your decision: why would you do what you just said?,’’` we find that in 96% of the cases they explicitly mention that they intend to put pressure on

[14]This exception is the Italy-Spain match in the quarter finals of the European Championship, June 2008. Gianluigi Buffon, the goalkeeper from Italy, won the toss and chose Spain to kick first.

the kicker of the second-kicking team, and that in no case they refer to the possibility of enhancing the performance of their own goalkeeper.

We interpret this evidence as supporting the hypothesis that subjects are perfectly aware of the psychological mechanism leading to pressure and underperformance. More importantly, they are not only aware, but respond optimally to it. This means, following the terminology used in behavioral economic theory, that they can be characterized as "sophisticates."

7 Concluding Remarks

Nature does not often create circumstances that allow a clear view of psychological principles at work. And, when it does, the phenomena are typically too complex to be empirically tractable in a way that we can discern psychological elements from within the complex behavior often exhibited by humans. Viewed from this perspective, the randomized experiment we have studied provides an unusual opportunity. Further, the setting involves highly incentivized professionals performing a simple, familiar task, in a real world strictly competitive situation.

The results provide support for a source of psychological pressure that has a detrimental effect on performance. The source that we identify is different from others, such as high stakes, social pressure or peer pressure. Since it is a source that is endogenous to the course of competition, the results seem relevant both from a theoretical and empirical perspective for competitive environments at large. They can also be taken to confirm the view that, much like cognitive abilities, emotional skills can have a major determining effect on performance and economic outcomes (Heckman, 2008).

Lastly, from the perspective of the recent behavioral economics literature, we find a significant and quantitatively important psychological effect not previously documented. From the perspective of rational choice theory, we find that individuals are aware of this effect and, when given the chance, they rationally respond to it.

APPENDIX A. (Not for Publication) Official FIFA Regulations

Away goals and extra time are methods of deciding the winning team where competition rules require one team to be declared the winner after a drawn match. A penalty shoot-out is held when a game remains tied after these methods have been applied. The procedure in force until July 2003 is described in the *Laws of the Game of FIFA* (2003, 2007), available in http://www.fifa.com, as follows:

"**Kicks from the Penalty Mark. Procedure:**

- The referee chooses the goal at which the kicks will be taken.
- The referee tosses a coin and the team whose captain wins the toss takes the first kick.
- The referee keeps a record of the kicks being taken.
- Subject to the conditions explained below, both teams take five kicks.
- The kicks are taken alternately by the teams.
- If, before both teams have taken five kicks, one has scored more goals than the other could score, even if it were to complete its five kicks, no more kicks are taken.
- If, after both teams have taken five kicks, both have scored the same number of goals, or have not scored any goals, kicks continue to be taken in the same order until one team has scored a goal more than the other from the same number of kicks.
- A goalkeeper who is injured while kicks are being taken from the penalty mark and is unable to continue as goalkeeper may be replaced by a named substitute provided his team has not used the maximum number of substitutes permitted under the competition rules.
- With the exception of the foregoing case, only players who are on the field of play at the end of the match, which includes extra time where appropriate, are allowed to take kicks from the penalty mark.
- Each kick is taken by a different player and all eligible players must take a kick before any player can take a second kick.
- Only the eligible players and match officials are permitted to remain on the field of play when kicks from the penalty mark are being taken.
- All players, except the player taking the kick and the two goalkeepers, must remain within the center circle.
- The goalkeeper who is the team-mate of the kicker must remain on the field of play, outside the penalty area in which the kicks are being taken, on the goal line where it meets the penalty area boundary line.
- Unless otherwise stated, the relevant Laws of the Game and International F.A. Board Decisions apply when kicks from the penalty mark are being taken.
- When a team finishes the match with a greater number of players than their opponents, they shall reduce their numbers to equate with that of their opponents

and inform the referee of the name and number of each player excluded. The team captain has this responsibility.

• Before the start of kicks from the penalty mark the referee shall ensure that only an equal number of players from each team remain within the center circle and they shall take the kicks.

After July 2003, the second point was replaced by:

• The referee tosses a coin and the team whose captain wins the toss decides whether to take the first or the second kick.

APPENDIX B. A Theoretical Model

This appendix studies the simplest model we can think of that is consistent with the empirical evidence, predicts a first-mover advantage and relies on a reference point associated with the partial score. After the presentation of the model, various extensions are discussed.

Let (s, r) denote the score $s \in Z$ at the end of round $r \geq 1$. The score measures the difference in goals between the team that kicks first F and the one that kicks second D. A round involves one penalty kick for F and one for D. The total number of rounds is n. The *partial score* for a team $\alpha \in \{F, D\}$ in round r is the difference between the goals scored by α and those scored by the opponent, immediately before team α is about to take its penalty kick in round r. That is, for team F the partial score at r is $(s, r-1)$, while, for team D, it is $(-s-x, r-1)$, where $x = 1$ if F scores in round r and $x = 0$ otherwise. In what follows we will use the terms team and player indistinctly.

Denote by $p \in [0, 1]$ the probability of player α scoring a goal when the partial score is tied or positive for him, and by $q \in [0, 1]$ the probability of player α scoring a goal when he is behind in the partial score by at least one goal. Under psychological pressure $p > q$, while under no psychological pressure $p = q$.

For any given (s, r) with $r < n$ there are exactly four possible outcomes at the end of round r: (i) both players score a goal, (ii) the first scores and the second fails, (iii) the first fails and the second scores, and (iv) both fail. The probability vectors associated to these outcomes depend on (s, r). There are three possible cases:

1. if $s = 0$, then $(p \cdot q, p(1-q), (1-p)p, (1-p)^2)$,
2. if $s > 0$, then $(p \cdot q, p(1-q), (1-p)q, (1-p)(1-q))$, and
3. if $s < 0$, then $(q \cdot p, q(1-p), (1-q)p, (1-q)(1-p))$.

To simplify notation we write $a = p \cdot q$, $b = p(1-q)$, $c = (1-p)p$, $d = (1-p)^2$, $e = (1-p)q$, and $f = (1-p)(1-q)$. The above defines a Markov chain. Since we are interested in rank-order tournaments, we need to refine the notion of maximum and minimum scores. If n is even, the maximum and minimum scores are $\frac{n}{2}+1$ and $-(\frac{n}{2}+1)$, while if n is odd the maximum and minimum scores are $\frac{n+1}{2}$ and $-(\frac{n+1}{2})$. The state space is formed by all possible scores $S = \{s_{max}, s_{max}-1, \ldots, -1, 0, 1, \ldots, s_{min}-1, s_{min}\}$ with s_{max} and s_{min} defined as above. Typical elements of S are denoted by s, t or $s_0, s_1, \ldots, s_n$. The transition matrix P follows from the single-step transition probabilities p_{st}:

$$p_{00} = a + d;\ p_{01} = b;\ p_{0,-1} = c$$

$$s \in \{s_{min}, s_{max}\},\ p_{ss} = 1$$

$$s \in S \setminus \{0, s_{min}, s_{max}\},\ p_{ss} = a + f$$

$$s \in S \setminus \{0, s_{min}, s_{max}\},\ p_{s,s+1} = p_{-s,-s-1} = b$$

$$s \in S \setminus \{0, s_{min}, s_{max}\},\ p_{s,s-1} = p_{-s,-s+1} = e$$

and $p_{st} = 0$ otherwise. The initial distribution μ puts all the probability mass in state 0. Denote by $T(n, P)$ the n-round sequential tournament between F and D with transition matrix P. Denote by $p_{st}^{(n)}$ the (s,t) entry in the n-th power of the transition matrix P. Since the Markov chain is stationary, $p_{st}^{(n)}$ represents the probability of reaching state t starting from state s in n rounds.

The probability that team F wins the n-round sequential tournament $T(n, P)$ is:

$$W(F, n) = \sum_{s=1}^{s=s_{max}} P(s, n),$$

with $P(s, n)$ denoting the probability of a final score s. To calculate $P(s, n)$ we have to correct for the probability of reaching a *final state* in some previous round. A final state is a pair (s, r) where there is no possibility of turning the sign of the score s around in the remaining time $n - r$. Then we have that

$$P(1, n) = p_{01}^{(n)} - p_{02}^{(n-1)} p_{21},$$

$$P(s_{max}, n) = p_{0,s_{max}-1}^{(n+1-s_{max})} p_{s_{max}-1,s_{max}} + p_{0s_{max}}^{(n+1-s_{max})},$$

and for $1 < s < s_{max}$

$$P(s, n) = p_{0,s-1}^{(n+1-s)} p_{s-1,s} + p_{0s}^{(n+1-s)} - p_{0,s+1}^{(n-s)} p_{s+1,s}.$$

In principle, these probabilities can be obtained using standard matrix algebra. The probability of team D winning at the end of the n-round contest $W(D, n)$ is obtained analogously.

Denote by $W(\alpha, r)$ the probability that α is either ahead of its opponent at the end of round $r \leq n$ or has already won the tournament by then. We are now ready to derive convenient formulations for $W(F, r)$ and $W(D, r)$.

Proposition 1 Let $T(n,P)$ be an n-round sequential tournament. Then, for every $r \leq n$, $W(F,r) = \frac{b}{b+c}(1-p_{00}^{(r)})$ and $W(D,r) = \frac{c}{b+c}(1-p_{00}^{(r)})$.

Proof of Proposition 1. Take any path ending in state $s > 0$ in round r, and denote it by $s_0 s_1 \cdots s_{r-1} s_r$ with $s_0 = 0$ and $s_r = s$. The probability measure of such path is $p_{s_0 s_1} \cdots p_{s_{r-1},s_r}$. We distinguish between two cases: the path reaches a final state s' in some previous round $h < r$, or the path does not reach such a final state s' in some previous round $h < r$.

Consider first the case where the path does not reach a final state in some previous round. We construct a unique symmetric path to the original one, ending in state $-s$. If $s_{r-1} = 0$, stop. Otherwise, proceed backwards until reaching a $0 \leq k \leq r-1$ such that $s_k = 0$. Clearly, such a k exists. Then, for every $l \geq k$ write $s'_l = -s_l$, and write $s'_l = s_l$ otherwise. It is immediate that the constructed path $s'_0 s'_1 \cdots s'_{r-1} s'_r$ starts with $s'_0 = 0$, ends in $s'_r = -s$, does not go through any final state, and has an associated probability measure of $p_{s'_0 s'_1} \cdots p_{s'_{r-1},s'_r}$, where $p_{s'_l,s'_{l+1}} = p_{s_l,s_{l+1}}$ for every $l \neq k$, while $b = p_{s_k,s_{k+1}} \geq p_{s'_k,s'_{k+1}} = c$. That is, the difference in the probability measures between the two paths is $(b-c)p_{s_0 s_1} \cdots p_{s_{k-1},s_k} p_{s_{k+1},s_{k+2}} \cdots p_{s_{r-1},s_r}$.

Consider now the case where the original path $s_0 s_1 \cdots s_{r-1} s_r$ does reach a final state s' in a previous round h. Firstly, we modify the path $s_0 s_1 \cdots s_{r-1} s_r$ to correct for the sub-path following the final state s', by writing $s_h = s_{h+1} = \cdots = s_{r-1} = s_r = s'$, with associated probability measure $\bar{p}_{s_l,s_{l+1}} = p_{s_l,s_{l+1}}$ whenever $s_l \leq h-1$, and $\bar{p}_{s_l,s_{l+1}} = 1$ whenever $s_l > h-1$. Secondly, apply exactly the same argument than above to the modified path, to show that there exists a unique symmetric path ending in final state $-s'$, and where the difference in the probability measures between the two paths is $(b-c)\bar{p}_{s_0 s_1} \cdots \bar{p}_{s_{k-1},s_k} \bar{p}_{s_{k+1},s_{k+2}} \cdots \bar{p}_{s_{r-1},s_r}$.

Consequently, it is immediate that there exists a probability mass $\gamma(r)$ such that $W(A,r) = b\gamma(r)$ and $W(B,r) = c\gamma(r)$. Note that by definition of final states, $p_{00}^{(r)}$ does not reach any final state in some previous round $h \leq r$. Now since $W(A,r) + W(B,r) + p_{00}^{(r)} = 1$, it follows that $\gamma(r) = \frac{1}{b+c}(1-p_{00}^{(r)})$, and hence $W(A,r) = \frac{b}{b+c}(1-p_{00}^{(r)})$ and $W(B,r) = \frac{c}{b+c}(1-p_{00}^{(r)})$.□

Since under no psychological pressure $p = q$ implies $b = c$, it directly follows from Proposition 1 that $W(F,n) = W(D,n)$. On the other hand, under psychological pressure, that is when $p > q$, we have that $b > c$, and hence Proposition 1 implies that $W(F,n) > W(D,n)$. We summarize the above in the following corollary:

Corollary 1 Let $T(n,P)$ be an n-round sequential tournament. Then, for every $p, q \in [0,1]$, and for every $r \leq n$,
- if $p = q$, $W(F,r) = W(D,r)$, and
- if $p > q$, $W(F,r) > W(D,r)$.

We now show that no other model with any possible heterogeneous configuration of probabilities of success that are contingent upon round, player and score can account for a first (or second) mover advantage. Denote the identity of a team by $k \in \{1,2\}$. A team k may go first in the tournament or second. Denote by $k(F)$ the case when player k goes first and by $k(D)$ when the same player k goes second. $\mathcal{P} = \{p_r^k(s)\}_{r \leq n, k \in \{1,2\}, s \in S}$ is the possibly heterogeneous collection of probabilities of success, and $w(k(\alpha), \mathcal{P}, n)$ is the probability of player k in the role of α winning the n-round tournament when $\mathcal{P}$. Thus, we claim that:

Lemma 1 For every $\mathcal{P}$ and n, $w(k(F), \mathcal{P}, n) = w(k(D), \mathcal{P}, n)$.

The proof of the lemma is immediate, and hence will be omitted.

<u>Discussion.</u> The main merit of the model is that it only takes two parameters to predict a greater probability of the first team winning. This simplicity is not without limitations, which further generalizations of the model may address. For instance, it is certainly possible to consider the difference $p - q$ to be team-specific, or simply a four-parameter model (p_i, q_i), $i = 1, 2$, to capture the possibility of order-dependent technologies. Clearly, for example, a partial score of 0 is better news for the second team than for the first team, and this may have an effect. These considerations will readily enrich the model and can be easily incorporated. In this case, it can be shown that the advantage of one team over the other depends on the parameter range of the (p_i, q_i) values, and that it is possible to find parameter values where both the first team has an advantage over the second and vice versa.

Second, the assumption that all team players use the same (p, q) technology can be generalized to introduce heterogeneity in player quality: $\{(p_i, q_i)\}_{i \in Q}$, where Q is the set of players ordered by quality. This opens up the possibility of investigating the strategic placement of players throughout the tournament, a decision that is to be taken before the toss out. We have studied this extension by having the two teams with (the same) three types of players differing in their (p, q) technology compete in a tournament with just three rounds. We found that, even in this simple case, it is possible to find parameters within an empirically sensible range of values that make each of the 27 possible combinations of the players a Nash equilibrium. That is, in theory, any combination could be optimal. This means that without precise information regarding the technologies of each player, there is no sharp, testable implication on the strategic allocation of players.

REFERENCES

Ariely, D., U. Gneezy, G. Loewenstein, and N. Mazar, "Large Stakes and Big Mistakes," *Review of Economic Studies,* forthcoming, 2008.

Arellano, Manuel and O. Bover. "Another Look at the Instrumental Variable Estimation of Error–Components Models," *Journal of Econometrics* 68:29–51, 1995.

Arellano, Manuel and Raquel Carrasco. "Binary Choice Panel Data Models with Predetermined Variables," *Journal of Econometrics,* 115:125-157, 2003.

Arellano, Manuel and Honoré, Bo E. "Panel Data Models: Some Recent Developments." In *Handbook of Econometrics*, vol. 5, chapter 53. Edited by J. E. Heckman and E. Leamer, North-Holland, 2001.

Aumann, Robert J., Foreword to 'Two-Sided Matching: A Study of Game-Theoretic Modeling and Analysis,' Econometric Society Monographs No. 18, Cambridge University Press, 1990.

Aumann, Robert J., "An Interview with Robert Aumann," *Macroeconomic Dynamics* 9, 683-740, 2005.

Becker, Gary S. *The Economic Approach to Human Behavior*, University of Chicago Press, 1976.

Becker, Gary S. *Accounting for Tastes.* Harvard University Press, Cambridge, MA, 1996.

Becker, Gary S. and Kevin M. Murphy. *Social Economics. Market Behavior in a Social Environment.* Harvard University Press, Cambridge, MA, 2000.

Beilock, Sian. "Choking under Pressure," in R. Baumeister and K. Vohs (Eds.) *Encyclopedia of Social Psychology.* Sage Publications, 2007.

Bhaskar, V. "Rational Adversaries? Evidence from Randomized Trials in One Day Cricket," *Economic Journal*, forthcoming, 2008.

Bronars, S. "Risk Taking in Tournaments," Mimeo, University of Texas, 1987.

Bull, C.; Schotter, A., and Weigelt, K. "Tournaments and Piece Rates: An Experimental Study." *Journal of Political Economy* 95:1-31, 1987.

Cabral, L. "Increasing Dominance with No Efficiency Effect," *Journal of Economic Theory* 102:471-479, 2002.

Cabral, L. "R&D Competition when Firms Choose Variance," *Journal of Economics & Management Strategy*, 12:139-150, 2003.

Cabral, L. and M. H. Riordan, "The Learning Curve, Market Dominance, and Predatory Pricing," *Econometrica* 62:1115-1140, 1994.

Caplin, A. S. and Leahy, J. "Psychological Expected Utility Theory and Anticipatory Feelings," *Quarterly Journal of Economics* 116:55-80, 2001.

Che, Yeon-Koo and Terry Hendershott, "How to Divide the Possession of a Football," *Economics Letters*, forthcoming, 2007.

Della Vigna, Stefano. "Psychology and Economics: Evidence from the Field," *Journal of Economic Literature,* forthcoming, 2007.

Ehrenberg, Ronald G., and Bognanno, Michael L. "Do Tournaments Have Incentive Effects?" *Journal of Political Economy* 98:1307-1324, 1990.

Epstein, Larry G. and Kopylov, Igor. "Cold Feet," *Theoretical Economics* 2:231-259, 2007.

Ericsson, K. A., N. Charness, P. J. Feltovich, and R. D. Hoffman, editors. *The Cambridge Handbook of Expertise and Expert Performance.* Cambridge University Press, 2006.

Fernandez-Val, I. "Fixed Effects Estimation of Structural Parameters and Marginal Effects in Panel Probit Models," Mimeo, Boston University, 2008.

FIFA, *Fédération Internationale de Football Association - Official Laws of the Game*, Zurich, Switzerland, 2003, 2007.

Gilbert, R. J. and Newbery, D. M. "Pre-emptive Patenting and the Persistence of Monopoly," *American Economic Review* 72:514-526, 1982.

Green, Jerry R., and Stokey, Nancy L. "A Comparison of Tournaments and Contracts." *Journal of Political Economy* 91:349-364, 1983.

Harrison, G. W., and J. A. List: "Field Experiments," *Journal of Economic Literature,* 42:1009-1055, 2004.

Heckman, James J. "Schools, Skills, and Synapses," IZA Discussion Paper No. 3515, 2008.

Honoré, Bo E. and Lewbel, Arthur. "Semiparametric Binary Choice Panel Data Models without Strictly Exogenous Regressors," *Econometrica* 70:2053-2063, 2002.

Honoré, Bo E. and Kyriazidou, Ekaterini. "Panel Data Discrete Choice Models with Lagged Dependent Variables," *Econometrica* 68:839–874, 2000.

Hume, David, *A Treatise of Human Nature.* New York: Oxford University Press. First edition, 1739 [1978]

Hvide, H. K. and E. G. Krinstiansen. "Risk Taking in Selection Contests," *Games and Economic Behavior* 42:172-179, 2003.

Hvide, H. K. "Tournament Rewards and Risk Taking," *Journal of Labor Economics* 20:877-898, 2002.

Kahneman, D. *Attention and Effort,* Englewood Cliffs, NJ: Prentice Hall, 1973.

Köszegi, Botond and Rabin, Matthew. "A Model of Reference-Dependent Preferences," *Quarterly Journal of Economics* 121:1133-1166, 2006.

Lazear, Edward P., and Rosen, Sherwin H. "Rank-Order Tournaments as Optimum Labor Contracts." *Journal of Political Economy* 89:841-864, 1981.

Loewenstein, G. "Anticipation and the Value of Delayed Consumption," *Economic Journal* 97:666-684, 1987.

Manski, Charles F., *Identification Problems in the Social Sciences*, Harvard University Press, 1995.

Morris, L., D. Davis and C. Hutchings, "Cognitive and Emotional Components of Anxiety: Literature Review and Revised Worry-Emotional Scale," *Journal of Educational Psychology* 73:541-55, 1981.

Nalebuff, B. and J. E. Stiglitz, "Prizes and Incentives: Towards a General Theory of Compensation and Competition," *Bell Journal of Economics* 14:21-43, 1983.

Palacios-Huerta, I. "Professionals Play Minimax," *Review of Economic Studies* 70:395-415, 2003.

Prendergast, Canice. "The Provision of Incentives in Firms." *Journal of Economic Literature* 37:7-63, 1999.

Rauh, Michael T. and Seccia, Giulio. "Anxiety and Performance: An Endogenous Learning-by-Doing Model," *International Economics Review* 47:583-609, 2006.

Rosen, Sherwin H. "Prizes and Incentives in Elimination Tournaments." *American Economic Review* 76:701-715, 1986.

Smith, Adam, *The Theory of Moral Sentiments.* London: Oxford University Press. First edition, 1759 [1976]

Wall Street Journal, "Should the Outcome of a Coin Flip Mean So Much in NFL Overtime?" December 22, 2003.

Table 1 – Description of the Dataset

Competition	Type	1970-2003 Shoot-outs *N*	1970-2003 Kicks *N*	2003-2008 Shoot-outs *N*	2003-2008 Kicks *N*	All Shoot-outs *N*	All Kicks *N*
World Cup	National Teams	16	153	4	33	20	186
European Championship	National Teams	9	97	4	42	13	139
American Cup	National Teams	12	116	3	31	15	147
African Nations Cup	National Teams	9	110	4	58	13	168
Gold Cup	National Teams	5	55	2	17	7	72
Asian Nations Cup	National Teams	--	--	7	74	7	74
Champions League	Clubs	8	82	12	117	20	202
UEFA Cup	Clubs	12	101	18	181	30	282
Spanish Cup	Clubs	29	308	26	259	55	567
German Cups	Clubs	24	273	44	479	68	752
English Cups	Clubs	5	48	9	94	14	142
All		129	1,343	133	1,385	262	2,731

* The dataset includes all the shoot-outs in the history of the World Cup, European Championship, American Cup, African Nations Cup (except one), and Gold Cup. All these are international competitions for national teams. The European Champions League and European UEFA Cup are international club competitions in Europe. For these two competitions the dataset includes all the shoot-outs that ever took place in the final match and all those that took place in any of the rounds in the period 2000-2008. The Spanish Cup, the German Cups and the English Cups are national club competitions. For the Spanish Cup and the German Cup and Supercup in Germany, the dataset has all the shoot-outs that took place in a final match, plus all the shoot-outs in all the rounds for the period 1999-2008 (Spanish Cup) and for the period 2001-2008 (German Cup). For the German Supercup it includes all those that ever took place. The English Cups include data on the F.A. Cup, League Cup and the F.A. Community Shield.

Table 2 – Pre-Treatment Characteristics

Criterion	*N*	First Team	Second Team	Difference
FIFA rankings	35	.46 (.50)	.54 (.50)	-.086 (1.011)
UEFA rankings	20	.37 (.49)	.63 (.49)	-.263 (.991)
Category	58	.50 (.35)	.50 (.35)	0.00 (.70)
Position (when same category)	30	.52 (.50)	.48 (.50)	.034 (1.017)
Experience	128	.48 (.32)	.52 (.32)	-.031 (.653)
Home team	82	.57	.43	.14

* FIFA publishes ranks on national teams since 1993. The UEFA ranking applies to international club team competitions (Champions League and UEFA Cup), and it is published since 1959. Teams taking part in national cup competitions--Spanish Cup, German Cups, and English Cups--may or may not belong to the same category or division in the national league competition. If they belong to the same one, we consider their standings in the league table at the time of the shoot-out. Experience refers to the number of previous shoot-outs in which a team has participated in which we observe in our dataset. The table reports the proportion in which each team shows a better entry in the respective criterion (=1 if higher, =0 if lower, =0.5 if same). Home team equals 1 if the team plays in its own stadium. Std. deviation in parentheses.

Table 3 – Determinants of Winning the Tournament

	Probit	Probit	Probit	Logit	Logit	Logit
Constant	-0.265**	-0.275	-0.263	-0.424**	-0.441	-0.421
	(0.111)	(0.197)	(0.436)	(0.180)	(0.320)	(0.708)
Team kicks first	**0.530*** **	**0.645*** **	**0.629*** **	**0.849*** **	**1.035*** **	**1.009*** **
	(0.158)	**(0.171)**	**(0.172)**	**(0.254)**	**(0.278)**	**(0.278)**
Home field		-0.087	-0.110		-0.1421	-0.178
		(0.208)	(0.209)		(0.337)	(0.338)
Neutral field		-0.043	-0.055		-0.070	-0.089
		(0.250)	(0.274)		(0.399)	(0.434)
Category (1 if higher)		-0.008	0.008		-0.012	0.014
		(0.171)	(0.172)		(0.277)	(0.278)
Home*Category			0.000			0.001
			(0.255)			(0.406)
Neutral*Category			0.001			0.004
			(0.072)			(0.186)
<u>Team kicks first interacted with:</u>						
*Home field			0.002			0.005
			(0.569)			(0.928)
*Neutral field			7.89e-05			-6.2e-06
			(0.508)			(0.817)
*Category			0.030			0.0001
			(0.536)			(0.876)
<u>Fixed effects for:</u>						
Champions League	No	No	Yes	No	No	Yes
UEFA Cup	No	No	Yes	No	No	Yes
National Teams	No	No	Yes	No	No	Yes
National Cups	No	No	Yes	No	No	Yes
N (teams)	258	224	224	258	224	224
Adjusted R^2	0.031	0.046	0.048	0.031	0.046	0.048
Log-Likelihood	-173.14	-148.10	-147.10	-173.14	-148.10	-147.10

Notes: *** and ** indicate significant at 1% and 5% respectively. Missing observations for “Category” for 17 shoot-outs.

Table 4 – Observations by Round and Winning Rates

Round	Number of Shoot-outs		If decided, percentage in which the	
Regular rounds 1-5	Observed	Decided	first team wins	second team wins
Round 1	128	0	-	-
Round 2	128	0	-	-
Round 3	128	1	100	0
Round 4	127	30	76.6	23.3
Round 5	97	63	60.3	39.7
	Total decided:	94	**65.9**	**34.1**
Decisive rounds				
Round 5*	42	20	60.0	40.0
Round 6	34	18	38.8	61.2
Round 7	16	4	75.0	25.0
Round 8	12	5	80.0	20.0
Round 9	7	1	0	100
Round 10	6	3	100	0
Round 11	3	1	0	100
Round 12	2	2	50.0	50.0
	Total decided:	54	**55.5**	**44.5**
All rounds	Total decided:	128	**60.5**	**39.5**

Notes: Rounds 5* are a subset of the Rounds 5 which began with the scored tied. That is, the outcome in these rounds is as decisive of the final outcome as it is that in rounds 6 and beyond. These rounds are counted twice, first within the Rounds 5 and second as “Decisive rounds.” We are missing the round by round data in one shoot-out.

Table 5 – Scoring Probabilities and Winning Frequencies by Team, Round and Partial Score

	First Team			Second Team		
	Behind	Even	Ahead	Behind	Even	Ahead
Round 1						
Scoring Probability	-	78.9	-	75.2	59.3	-
N	-	128	-	101	27	-
% Win Shoot-out	-	60.2	-	32.7	66.7	-
Round 2						
Scoring Probability	100	74.7	96.0	82.2	65.8	-
N	16	87	25	90	38	0
% Win Shoot-out	31.3	57.5	88.0	32.2	57.9	-
Round 3						
Scoring Probability	80.0	76.8	76.5	63.2	69.4	40.0
N	25	69	34	87	36	5
% Win Shoot-out	24.0	59.4	88.2	23	72.2	100
Round 4						
Scoring Probability	76.7	71.7	75.0	66.2	69.4	77.8
N	30	53	44	71	36	9
% Win Shoot-out	13.3	62.3	88.6	21.1	75.0	100
Round 5						
Scoring Probability	74.1	76.2	71.4	62.5	70.0	-
N	27	42	28	40	30	-
% Win Shoot-out	14.8	52.4	96.4	30.0	83.3	-
Rounds 6+						
Scoring Probability	-	67.5	-	68.5	65.4	-
N	-	80	-	54	26	-
% Win Shoot-out	-	58.8	-	24.1	76.9	-

Notes: The "scoring probability" for each team-score-round case is computed as the percentage of teams that scored a goal in that score-round situation. The "% Win Shoot-out" is the percentage of teams observed at a given score and round that *eventually* won the shoot-out. The number of observations *N* in rounds 6+ for each partial score and type of team (first or second) is computed as the sum of the number of teams that in rounds 6 and beyond are observed at a given partial score. That is, since the first team can be observed in various rounds with an even partial score and the second team can be observed in various rounds with the same or different (behind and even) score, the same team may be observed at multiple occasions. The scoring probabilities and the percentage of teams that win the shoot-out are computed using these as the number of observations. Round-by-round data is available upon request. We are missing the round by round data in one shoot-out

Table 6 – Random Effects Dynamic Panel Data Model

	[1]	[2]	[3]
Constant	**0.95*** **	**0.89*** **	**0.83*** **
	(6.77)	(3.22)	(2.88)
Round 2	0.87	0.23	0.12
	(0.23)	(0.11)	(0.63)
Round 3	0.02	0.45	0.00
	(0.45)	(1.12)	(0.50)
Round 4	1.27	1.20	1.21
	(1.11)	(1.33)	0.91)
Round 5	1.35	0.63	0.31
	(1.23)	(0.86)	(0.87)
Rounds 6+	-1.55	-1.37	-1.21
	(1.57)	(1.40)	(1.51)
Home field	-0.27	1.77	2.01
	(0.00)	(0.34)	(1.07)
Neutral field	0.03	0.35	0.66
	(0.06)	(1.01)	1.27)
International Competition	-0,03	0.20	-0,02
	(0.38)	(0.18)	(0.03)
Final match	-1.20	**-0.07***	**-0.02***
	(0.99)	(1.70)	(1.81)
Lagged penalty outcome	-0.27	-0.35	0.01
	(1.02)	(0.80)	(1.30)
Partial Score +1	0.98	1.21	2.17
	(1.23)	(1.00)	(0.88)
Partial Score 0	2.34	1.82	1.55
	(0.45)	(0.25)	(0.67)
Partial Score -1	**-0.21*** **	**-0.16*** **	**-0.13*** **
	(5.21)	(4.28)	(3.66)
Kicking Second		**-0.07** **	-0.03
		(1.97)	(1.60)
Partial Score -1 x Kicking second		**0.04** **	**0.02** **
		(2.32)	(1.98)
Partial Score -1 x Round 6+		**0.03** **	0.03
		(2.25)	(1.44)
Goalkeeper Fixed Effects	No	No	Yes
Competition Type Fixed Effects	No	No	Yes

Notes: The endogenous variable equals 1 if the penalty is scored and 0 otherwise. The sample has 1,343 individual penalty kick observations. Absolute, choice–biased sampling, heteroskedastic–consistent, *t*–statistics are reported in parentheses. The model is estimated by GMM. ***, **and * indicate significant at 1%, 5% and 10% respectively.

Table 7 – Marginal Effects

These marginal effects represent the percentage change in the probability of scoring conditional on each transition among states *i* to *j*. In Panel A, the states are partial scores; the current state is represented by *i* =0 at a given round, and *j* = +1, -1. In Panel B, the state is the kicking order and current state is represented by kicking second. These effects are computed from specification [3] in Table 6.

Panel A: Partial Scores

	Round 1	Round2	Round 3	Round 4	Round 5	Rounds 6+
From 0 to +1						
Kick first	-	1.42%	1.62%	1.36%	3.71%	-
Kick second	-	2.37%	2.51%	3.52%	-	-
From 0 to -1						
Kick first	-	-1.82%	-2.10%	-3.92%	-5.94%	-
Kick second	-3.01%	-0.42%	-1.36%	-2.33%	-4.55%	-6.97%

Panel B: Kicking First instead of Kicking Second

Round 1	Round 2	Round 3	Round 4	Round 5	Rounds 6+
0.28%	0.34%	0.56%	1.03%	1.40%	2.52%

Table 8- Determinants of Misses and Saves

Panel A: Proportions of Penalties Scored, Saved and Missed

	N (penalties)	Scored	No Scored: Saved	No Scored: Missed
First team	234	75.6%	17.1%	7.3%
Second team	220	68.6%	16.8%	14.6%
	Difference:	**-7.0%**	-0.3%	**+7.3%**

Panel B: Panel Data Analysis for Misses and Saves

	Misses [1]	Misses [2]	Saves [3]	Saves [4]
Constant	-1.61***	-1.77***	-0.92**	-0.73**
	(3.15)	(2.88)	(2.35)	(2.20)
Partial Score -1	**0.10*** **	**0.19*** **	-0.07	-0.08
	(3.22)	(3.01)	(0.55)	(0.26)
Partial Score 0	0.34	0.27	0.05	0.04
	(0.34)	(0.32)	(0.07)	(0.06)
Partial Score +1	0.19	0.21	0.96	-0.37
	(1.03)	(0.97)	(0.28)	(0.26)
Kicking Second	0.17	0.05	0.11	-0.34
	(0.77)	(0.89)	(1.26)	(1.25)
Partial Score -1 x Kicking second	-0.012**	-0.009**	0.008**	0.02*
	(2.33)	(2.32)	(1.17)	(0.85)
Field fixed effect?	Yes	Yes	Yes	Yes
Competition Type fixed effect?	Yes	Yes	Yes	Yes
Match Type fixed effect?	Yes	Yes	Yes	Yes
Round Number fixed effect?	No	Yes	No	Yes
Goalkeeper fixed effect?	No	Yes	No	Yes
N	454	454	454	454

Notes: Misses by kicker are penalty kicks shot to the upright posts, the horizontal crossbar or outside the goal. Saves are penalty kicks stopped by the goalkeeper. Absolute, choice–biased sampling, heteroskedastic-consistent, *t*–statistics are reported in parentheses. ***, **and * indicate significant at 1%, 5% and 10% respectively.

Table 9 – Survey Results

The following two questions were asked to soccer coaches and players:

Q1: "Assume you are playing a penalty shoot-out. You win the coin toss and have to choose whether to kick first or second. What would you choose: first; second; either one, I am indifferent; or, it depends?"

		Proportion answering:			
	N	First	Second	Indifferent	Depends
Coaches:					
Professional	21	90.5%	0	0	9.5%
Amateur	37	94.6%	0	0	5.4%
Players:					
Professional	67	97.0%	0	1.5%	1.5%
Amateur	117	96.5%	0	2.5%	1.0%
All :	242	**95.9%**	**0.0%**	**1.6%**	**2.5 %**

Q2: "Please explain your decision. Why would you do what you just said?"

Notes: Professional coaches and players come from the professional leagues in Spain (Primera Division and 2A and 2B Divisions). Amateur coaches and players come from Division 3 and Regional Leagues in Spain. The 4 coaches that answered "Depends," further explained that they would let their players choose what they preferred to do.

Figure 1 – Winning Frequencies in Aggregate Data

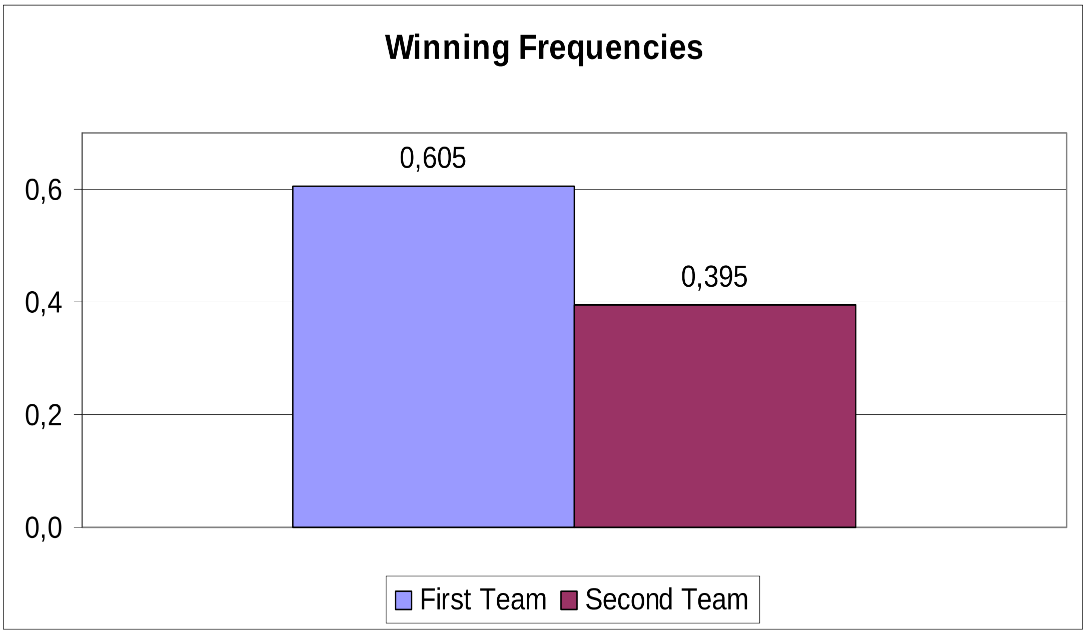


Figure 2.1

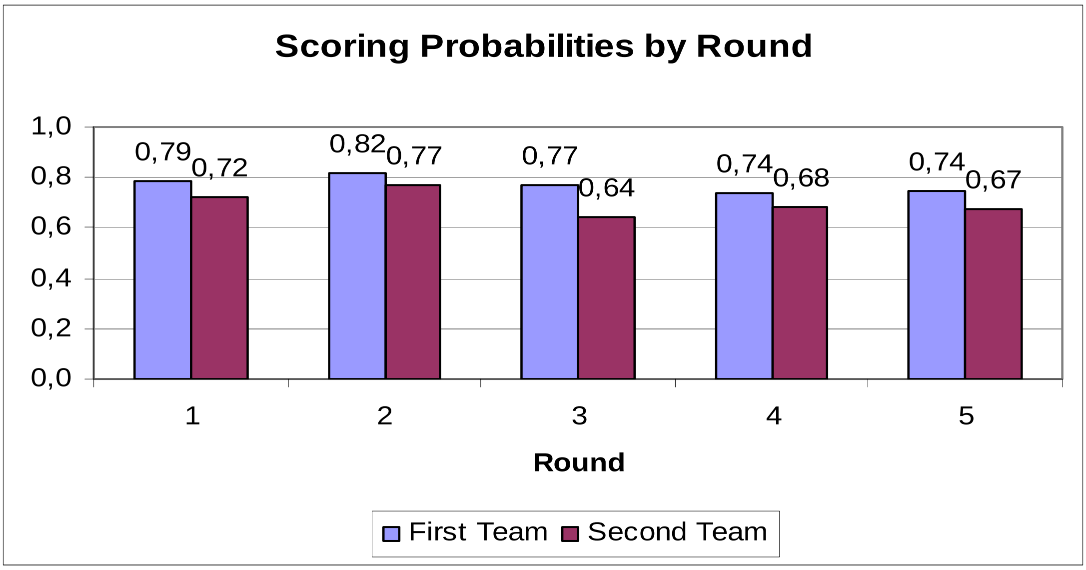
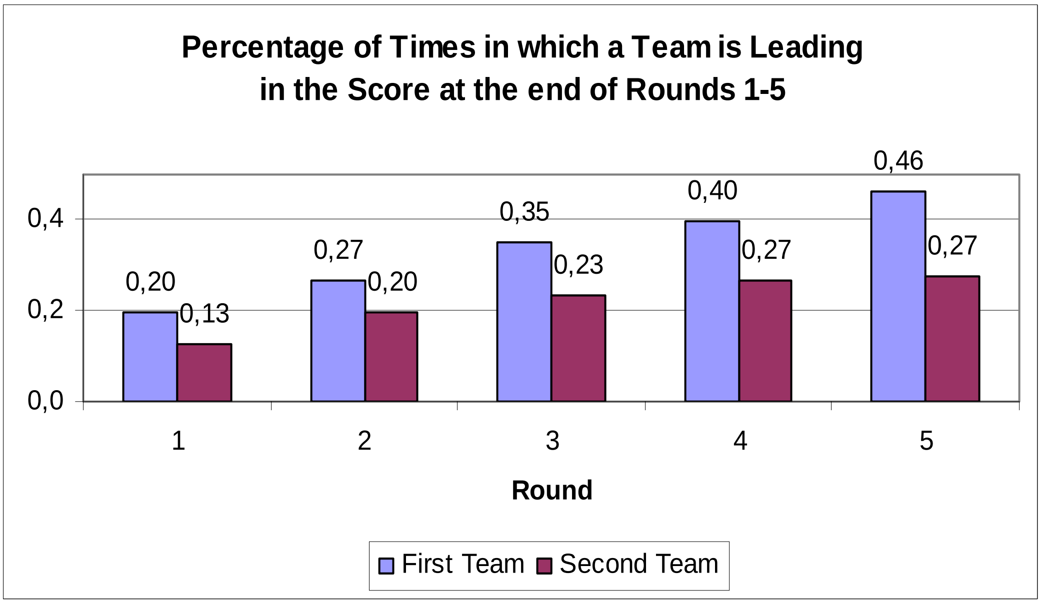


Figure 2.2



Note: If before both teams have taken five penalty kicks, one has score more goals than the other could possibly score even if it were to complete its five kicks, no more kicks are taken. The percentage of times in which a team is leading in the score at the end of a round in Figure 2.2 includes these cases; that is, cases in which the shoot-out already ended before this round, whereas in Figure 2.1 the scoring rate is only computed for the teams that are observed to kick in the corresponding round.