5

PSYCHOLOGICAL PRESSURE ON THE FIELD AND ELSEWHERE

The secret of getting ahead is getting started.
—Attributed to Mark Twain (1835–1910)

In the early 1960s, Rafael Ballester was a prestigious journalist in Cádiz, a city in the south of Spain. Cádiz is well known in the world of football because for decades it has organized a famous international summer tournament in early August, the Trofeo Ramón de Carranza. Each year, four different teams are invited. They play four games in two days, the semifinals on Saturday and, on Sunday, the match that determines the third and fourth place in the tournament is played, followed by the final.

Quite often the semifinals on Saturday ended up being tied, and the teams had to play for extra time. If they remained tied after the extra time, then, in keeping with tradition, a coin would decide the team that would play in the final the following day. This was the prevailing system in FIFA to break ties up to 1970. It was problematic. Teams would often be quite tired after the additional effort from the previous day and, with no rest days, the quality of soccer would suffer substantially on Sunday. Plus, there was, of course, the added unpleasant feature that many times it was the arbitrariness of a coin toss that decided the outcome of a match, sometimes even the final winner of the tournament.

Mr. Ballester had an ingenious idea to overcome these problems: ties would be resolved with a *penalty shoot-out* where both teams would kick the same number of penalty kicks until one has scored more goals than the other could score. For instance, both teams could begin by kicking five penalty kicks each. If they remained tied, they would kick more, until a winner was declared.

He published his idea in the newspaper *Diario de Cádiz* in August 1957, right after the final between Athletic Club de Bilbao and Sevilla that was decided by a coin.[1] The organizers liked it and decided to adopt it. The first opportunity to put the new system into action was on September 2, 1962, in the final match of the tournament between Barcelona and Zaragoza. Barcelona won after six penalty kicks. This procedure to break ties quickly gained popularity and spread to several friendly tournaments in Europe, Africa, and South America in following years. Soccer connoisseurs may remember, for instance, the Trofeo Corpus Christi from 1964, played in Orense (Spain), which featured three teams: FC Porto from Portugal, RC Deportivo La Coruña from Spain, and Athletic Club de Bilbao. The first game between RC Deportivo La Coruña and Athletic Club de Bilbao ended 1–1. In the penalty shoot-out, first Deportivo kicked five penalty kicks in one go. José Ángel Iribar, one of the best goalkeepers in soccer history, stopped four of them (an incredible performance since around 80% of penalties are scored on average), and the fifth one hit the goalpost. Then Athletic Club de Bilbao scored just its first penalty kick and won the match.

The popularity of the new system to break ties was such that, in 1970, FIFA decided to adopt it. There are no detailed minutes of the International Board Meeting held on June 27, 1970, at the Caledonian Hotel in Inverness, Scotland, when the shoot-out proposal was approved, but the idea of one team taking all penalties in one go was replaced by the system of alternate penalties that we know today.

Beginning on that date, the method of determining the winning team, where competition rules require that one team is declared the winner after a drawn match, was by a penalty shoot-out. Although it was too late for the Mexico World Cup in 1970, this decision meant that it would be 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 then on.

1 The story is reported in Relaño (2010). Some sources mistakenly credit Israeli Yosef Dagan as the inventor of the penalty shoot-out. After watching the Israeli team lose an Olympic quarter-final by drawing of lots in 1968, he proposed this system in a letter to the Israel Football Association. Others credit former German referee Karl Wald in a proposal to the Bavarian Football Association in 1970.

The basic rules of a penalty shoot-out 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.

History says that the first penalty shoot-out in a senior official competitive football match took place in England on August 5, 1970, just a few days after the FIFA approval. Manchester United was the first winner, defeating Hull City 4–3 on penalties in the semifinal of the Watney Cup. The set of five players from the first team that kicked in a penalty shoot-out included some of the greatest players ever: George Best, Brian Kidd, Bobby Charlton, Denis Law, and Willie Morgan, and Alex Stepney was in goal. First trivia alert: The first player to score in a shoot-out was George Best, with a low shot to the keeper's right. Second trivia alert: The first player to miss a kick in a shoot-out was Denis Law. Hull's keeper, McKechnie, dived to his right to save it. And a trivia question: What was the world's longest penalty shoot-out? The answer is in chapter 10.

This system was in place until July 2003, when FIFA decided to change the first regulation in the procedure slightly by replacing it with the following:

- 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 discussed in the previous chapters, offer substantial advantages for studying the role that psychological elements (emotions) may play in dynamic competitive environments. As Miller (1998) notes, right from the beginning, the *Daily Telegraph* confirmed the presence of emotional elements in this setting. After the Manchester United–Hull final, it wrote: "This was the first time this method of settling a match had been used at senior level in England and it must be rated a resounding success. The suspense, as five players from each side fired alternately, was almost intolerable." The

Palacios-Huerta_1st-proof.indd 70 2/11/14 11:37 AM

Daily Mail said, "The penalty-taking session which settled this pulsating game was one of the most exciting and dramatic features I have ever seen on a soccer field."

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 rationality paradigm. This idea suggests that any rational theory of human behavior that omits these elements may yield results of unknown reliability until confronted with the data.

Motivated by evidence from new and richer data sets during the past couple of decades, an important body of research has attempted to parsimoniously incorporate psychological motives into standard economic models. The empirical evidence to test these models is typically obtained from the observation of human decision-making in laboratory environments, where experiments have the important advantage of providing control over relevant margins. A great deal of laboratory evidence has been accumulated demonstrating circumstances under which strict rationality considerations break down and other patterns of behavior, including psychological considerations, emerge. Nature, however, is less willing to contribute with empirical evidence. In fact, it rarely creates the circumstances that allow a clear view of the psychological principles at work. And when it does, the phenomena are typically too complex to be empirically tractable in a way that allows psychological elements to be discerned within the characteristically complex behavior exhibited by humans.[2]

This is why a penalty shoot-out is important. It provides an unusually clean opportunity in a real-world environment to discern the presence of psychological elements. In addition to the virtues of a penalty kick described in previous chapters,

1. A penalty shoot-out is a randomized natural experiment, that is, a real-life situation in which the treatment and control groups are determined via explicit randomization. In this case, the treatment that is randomly given to one team is the *order* of play: One team goes first in the sequence of tasks (penalty kicks) and the other second. As is well known, randomized experiments provide researchers with the critical advantage that they guarantee that the conditions for causal inference are satisfied.
2. The subjects involved in a shoot-out are professionals who have to perform a simple task: kick a ball once.

2 See Rabin (1998) and DellaVigna (2009) for excellent surveys of existing work.

3. All the relevant variables that are typically hard to observe and measure in other settings can be observed and measured.
4. And, finally, the analysis of a penalty shoot-out is also important scientifically because it relates to several strands of literature in economics and psychology:

   a. First, the natural setting corresponds to what is known as a tournament. Tournament competitions are pervasive in organizations and in real life and often characterize situations such as competitions for promotion in internal labor markets in firms and organizations, patent races, political elections, and many others. As a framework of analysis, the tournament model was formally introduced by Lazear and Rosen (1981), and over the past couple of decades a large literature has studied a number of important aspects of this incentive scheme both theoretically and empirically.[3]

      Despite the large body of work, however, there is very little evidence documenting how psychological or emotional effects may be relevant in explaining the performance of subjects competing in tournament settings. Difficulties in clearly observing actions, outcomes, choices of risky strategies, and other relevant variables in a real-life tournament are often exceedingly high, and as a result it is typically impossible to discern the extent to which psychological elements may explain performance with sufficient precision.

      The characteristics of a penalty shoot-out, however, are ideal for overcoming these obstacles. Variables such as the choice of effort levels and risky strategies that are typically hard to observe and measure play no role in this setting: The task (kicking a ball once) involves little physical effort and, with only two possible outcomes (score or no score), risk plays no role either.[4] Outcomes (goal or no goal) can be perfectly observed and are immediately determined after players make their choices. The fact that there is no subsequent play and that the task is immediate (a penalty kick takes less than half a second) is indeed critical to cleanly interpreting the empirical evidence.

3 See Nalebuff and Stiglitz (1983) and Rosen (1986) for early contributions, and Prendergast (1999) for a review.

4 The role of risk in tournament competitions has been studied in Hvide (2002) and Hvide and Kristiansen (2003). In dynamic competition games, there is a literature on the “increasing dominance” effect of a leader over a rival, which studies the strategic amount of resources to use and allocate throughout a competition (Cabral 2003).

b. Second, 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 other aspects.[5] In a penalty shoot-out, however, both teams have the same stakes and both perform in front of the same audience. The explicit randomization procedure that is used to determine the kicking order means that there is no reason why one team should be systematically more affected than the other team by the stakes or the audience. The coin does not know which team is supported by the home audience (if any) or has greater stakes.

   What is new from the perspective of the existing academic literature is that differences in the interim state of the competition caused by the randomly determined kicking order may generate differences in psychological elements that could have an effect on performance.[6]

c. Finally, 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 effect 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 (2008) for the case of extra-time sudden-death regulations in the US National Football League). In a penalty shoot-out setting, however, we are under ideal circumstances: A coin flip determines only the order of competition, and both teams have exactly the *same* chances to perform a task. Yet, human nature may be such that the outcome of a perfect randomized trial has to be considered ex post unfair if in fact the order is shown to matter for performance.

We take data from the Union of European Football Associations (UEFA), the Rec.Sport.Soccer Statistics Foundation, the Association of Football Statisticians in the United Kingdom, the Spanish newspapers *Marca* and *El Mundo Deportivo,* www.weltfussball.de, and the archives of various soccer clubs. The data set comprises 1,001 penalty shoot-outs with 10,431 penalty kicks over the period 1970–2013. It is comprehensive in that it includes virtually all the penalty shoot-outs in the history of the

5 See, for instance, Ericsson et al. (2006) and Beilock (2010). Ariely et al. (2009) review and discuss this literature.

6 In contrast to the size of the psychology literature, the economics literature is fairly limited, with pioneering theoretical contributions by Loewenstein (1987), Caplin and Leahy (2001), and Rauh and Seccia (2006) on anxiety and anticipatory emotions. There are, however, no previous empirical contributions with evidence from strictly competitive environments in real life.

main international elimination tournaments involving national teams (e.g., the World Cup, European Championship, and American Cup) and club teams such as UEFA Champions League and the UEFA Cup (now known as the Europa League). It also includes data on national club elimination tournaments such as the Spanish Cup, German Cup, and the English Football Association Cup.

This chapter follows Apesteguia and Palacios-Huerta (2010), AP henceforth, and for every shoot-out of every competition, it collects information on the date, the identity of the teams kicking first and second, the final outcome of the shoot-out, the outcomes of each of the kicks in the sequence, 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) and variables that measure the quality of the teams, such as their previous experience in shoot-outs, their official FIFA and UEFA rankings (for national teams), and the division, category, and standings (for club teams).[7]

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 the following:

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

Next, we extend the analysis in AP (2010) by studying not only the data for 1970–2003 but also the data for the following decade as well, that is, 43 years: 1970–2013. Note that the average treatment effect is identical before and after 2003. The fact that after 2003 players are required to choose the order (whether to kick first or second) is irrelevant for the

7 Consistent with the randomization procedure used to determine the order of play, it is not possible to reject the null hypothesis that any of these characteristics are irrelevant in determining the order of play, at the usual levels of significance. That is, the coin does not systematically select a specific type of teams with certain characteristics to kick first or second.

size of the average treatment effect. Their choices are interesting as a test of rationality, or consistency, but it does not affect $T(B|x)$.

To see this effect, consider a shoot-out between teams i and j in the framework of Bhaskar (2009). Let w denote the state of the world that captures all relevant factors including the characteristics of the two teams, and let $p(w)$ be the win probability for i when i shoots first, and $q(w)$ the win probability when it shoots second. Under random assignment of the treatment "shooting first" (period 1970–2003), the probability that the team that shoots first wins is given by $0.5\{p(w) + [1 - q(w)]\} = 0.5[1 + \lambda(w)]$ where $\lambda(w) = p(w) - q(w)$. Obviously, $\lambda(w)$ can be negative for some w. Let us call this number $E(\lambda)$. Consider now the period after 2003, where the winner of the coin toss chooses the order. If the players always choose optimally, then the win probability for the team kicking first is exactly identical to $E(\lambda)$. But consider the opposite scenario: The winner of the coin toss always makes the inferior choice, that is, the winner chooses first when it should choose second, and second when it should choose first. Then the estimated treatment effect is also *exactly equal* to $E(\lambda)$. And the same, of course, in any intermediate scenario where the winner sometimes chooses to kick first and others second. All we can conclude after 2003 is the rationality or irrationality (the correctness or incorrectness) of the choices the teams make; the average treatment effect remains unchanged.[8]

Figure 5.1 and table 5.1 report $T(B)$ unconditional on any variables. The data show that kicking first conveys a strongly significant (beyond the 1% level) and sizable advantage: The team that kicks first wins the penalty shoot-out around 60% of the time.

Thus, the data show that a penalty shoot-out is not a 50–50 lottery. It is more like a 60–40 lottery where the first-kicking team has 20% more tickets. As expected, using a regression framework to provide an estimate of the treatment effect conditional on the complete set of available characteristics for the teams under various probit and logit specifications yields the same results. The order of play is strongly significant in every specification in table 5.2, and there is a significant and sizable advantage to the team that is first to kick. Mapping the regression coefficient into the corresponding Normal and Logistic distribution yields an effect in the most complete specifications of columns two and four in table 5.1, again, slightly above 60% for the team that kicks first.

What the clean natural experiment just studied allows us to identify is that the nature of the mechanism generating these differences in performance is psychological. These emotional effects are endogenous to

8 Bhaskar (2009) offers a more detailed analysis, with an excellent application to the consistency of batting choices in cricket.

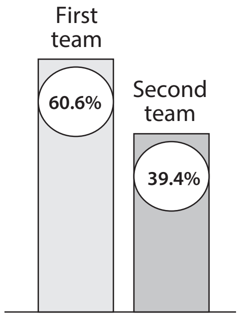

Figure 5.1. Winning frequencies by team, 1970–2013.

Table 5.1. Percentage First Team Wins in International and National Competitions 1970–2012

	Number of shoot-outs	First team wins (%)	
International Competitions			
1. National Teams			
World Cup	22	59.1%	
European Championship	15	33.3%	
Copa América	18	61.1%	
African Nations Cup	20	60.0%	
Gold Cup	10	70.0%	
Asian Cup	16	56.3%	
2. Club Teams			
European Champions League	49	63.3%	
European Cup Winners' Cup	32	62.5%	
UEFA Cup	110	55.5%	
National Competitions			
German Cups	183	49.7%	
English Cups	179	53.6%	
Spanish Cup	347	72.3%	
All International Competitions	292	57.8%	*p*-value: 0.0139
All National Competitions	709	61.0%	*p*-value: <0.0001
Total	1001	60.6%	*p*-value: <0.0001

Palacios-Huerta_1st-proof.indd 76 2/11/14 11:37 AM

Table 5.2. Determinants of Winner of Penalty Shoot-Out

	Probit	Probit	Logit	Logit
Constant	−0.267	−0.273	−0.437	−0.403
	(0.217)	(0.506)	(0.343)	(0.609)
Team kicks first	0.657***	0.633***	1.027***	1.012***
	(0.140)	(0.134)	(0.192)	(0.187)
Home field	−0.092	−0.114	−0.128	−0.165
	(0.210)	(0.244)	(0.352)	(0.340)
Neutral field	−0.052	−0.048	−0.073	−0.079
	(0.275)	(0.314)	(0.422)	(0.412)
Category	0.002	−0.007	0.011	−0.007
(1 if higher)	(0.182)	(0.170)	(0.272)	(0.228)
"Team kicks first" interacted with				
Home field	No	Yes	No	Yes
Neutral field	No	Yes	No	Yes
Category	No	Yes	No	Yes
N (teams)	2002	2002	2002	2002
Adjusted R^2	0.106	0.108	0.106	0.108

Note: Regressions in columns 2 and 4 also include fixed effects for Champions League, UEFA Cup, National Team, and National Cup competitions, as well as interactions between Home and Neutral field and Category.

the state of the competition itself and contribute to determining human performance in a strictly competitive (zero-sum) setting. What is not possible to identify, however, is the precise psychological mechanism that generates the result. We may speculate that the randomly determined order could generate differences in arousal, in anxiety, in shifting of mental process from "automatic" to "controlled," or in the narrowing of attention. Maybe it also generates differences in reference points. Köszegi and Rabin (2006), for instance, develop a model where a person's reference point is her or his rational expectation of the outcomes and "gain–loss" utility evaluations around this point influence her or his behavior. In a penalty shoot-out, the score at the time a player has to perform his or her task (the "ahead–behind" asymmetry caused by the order) may perhaps act as a reference point that has an effect on behavior.

Although we cannot answer the question of what is the specific psychological mechanism at play in this effect in performance, we can attempt to answer other questions:

1. Are subjects aware of the advantage of going first?
2. Do they rationally respond to this advantage by systematically choosing to kick first when given the choice (after 2003)?
3. Do players talk about a specific psychological mechanism that is at work in generating these effects?

According to a survey conducted in AP (2010), the answer to the first two questions is affirmative (see table 5.3).

Clearly, if subjects are aware of the effect, they should always choose to go first. Unfortunately, there are no public records of players' choices because FIFA regulations do not require referees to record this information. By watching matches that end in a penalty shoot-out, it is sometimes possible (when the TV channel is not airing commercials), to catch the instant when the referee flips the coin and talks to the winner of the toss. Consistent with their answers in the survey, in every case when it was possible to see the coin toss, the winner of the toss was observed to choose to kick first, with just two exceptions. The first exception is the Italy–Spain match in the quarter-finals of the European Championship in June 2008. Gianluigi Buffon, the goalkeeper from Italy, won the toss against Iker Casillas, the goalkeeper from Spain, and chose Spain to kick first. Interestingly enough, the second exception involves the

Table 5.3. Survey

	Observations	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	1.6%	2.5%

Notes: The following 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?"

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

Professional coaches and players come from the professional leagues in Spain (Primera División and 2A and 2B División). Amateur coaches and players come from División 3 and regional leagues in Spain. The four coaches who answered "It depends" further explained that they would let their players choose what they preferred to do.

Palacios-Huerta_1st-proof.indd 78 2/11/14 11:37 AM

same teams and the same players five years later. In the semifinal of the Confederations Cup in June 2013, Casillas won the toss this time and decided to return the favor: He chose Italy to kick first. Perhaps goalkeepers are, after all, different from other players. Or as the old saying goes, you do not have to be crazy to be a goalkeeper, but it helps.

Finally, with regard to the third question, most subjects argue that their choice is motivated by the desire to put pressure on the kicker of the opposing team. Coding their answers to this question, in 96% of the cases they explicitly mention that they intend to put psychological pressure on the second kicking team, something which is consistent with the evidence reported in AP (2010) that kickers decrease their performance when lagging (as opposed to goalkeepers, who improve theirs when leading).

*

A main difficulty for identifying the specific psychological mechanism at play is that a penalty kick involves two people, not one, and so the effect could arise from one player, from the other, or from both. An idea then is to look at similar sports settings that involve analogous dynamic decision-making processes but involve just one individual, not two, and also to look at other competitive activities with two individuals.

Mertel (2011) looks at data on more than 220,000 free throws from four seasons of professional NBA basketball. Carefully controlling for reverse causality, serial correlation, and a number of potential factors, he finds that players are significantly *more* likely to hit their free throws when they are *ahead* on the scoreboard than when they are behind. The difference in the scoreboard stops being relevant once the outcome of the game is beyond doubt and players revert to their inherent ability-reflecting mean. These findings are important because a free throw is an individual nonstrategic task and the results are consistent with the evidence from penalty shoot-outs: A leading or lagging asymmetry in a dynamic competition causes differences in performance.

In a golf setting, Pope and Schweitzer (2011) analyze more than 2.5 million putts in tournaments of the PGA Tour using precise laser measurements. They find that even the best golfers—including Tiger Woods—show evidence of loss aversion (Kahneman and Tversky 1979): Professional golfers hit birdie putts less accurately than they hit otherwise similar par putts. Golf provides a natural setting to test for loss aversion because golfers are rewarded for the total number of strokes they take during a tournament, and yet each individual hole has a salient reference point, par. When hitting a birdie, a player is "leading" over the hole, whereas when hitting a par, the player is "lagging" and has a chance to "tie" the hole. As indicated already, Köszegi and Rabin (2006) model a person's reference point as her or his expectations about outcomes, and gain–loss utility evaluations around this point influence her or his behavior. In golf, par seems a natural reference point, and in a penalty shoot-out it is possible to conjecture that the score at the time a player kicks acts as a reference point. Consistent with this reference-point hypothesis, the accuracy gap between par and birdie putts diminishes for very difficult holes and the gap between par and bogey putts widens for very difficult holes. A difficulty in this golf setting, however, is that risk taking and performance cannot be measured separately.

Perhaps the cleanest evidence showing that an interim rank (a leading–lagging asymmetry) in a dynamic competition affects performance comes from weightlifting, which, like a free throw in basketball, is an individual, nonstrategic task. Genakos and Pagliero (2012) empirically study the effect of interim rank on performance using data on professionals competing in tournaments for large rewards. The fact that risk plays a role in this setting would appear to make the empirical identification difficult. However, the authors observe both the intended action (competitors announce the weight they want to lift) and the performance of each participant, and so they can measure risk taking and performance separately. They obtain two important findings. First, risk-taking exhibits an inverted-U relationship with interim rank. Revealing information on relative performance induces individuals trailing just behind the interim leaders to take greater risks. Second, and most relevant in the context of this chapter, competitors systematically *underperform* when ranked closer to the top, despite higher incentives to perform well. In other words, disclosing information on relative ranking *hinders* interim leaders.

Although the identification of the exact channel through which emotions affect performance remains an open question, these different results from other sports on nonstrategic tasks are consistent with the hypothesis that information on relative performance hampers performance by increasing psychological pressure when subjects are lagging in the competition.

The implications of this phenomenon may be wide ranging and perhaps extend to other areas. For instance, Heckman (2008) remarks that emotional skills help determine a number of socioeconomic outcomes, contribute to performance at large, and even help to determine *cognitive* achievement. Understanding whether or not psychological elements that determine performance in noncognitive tasks (kicking a soccer ball, weightlifting, golf, basketball) may also contribute to explaining cognitive performance is a fascinating issue. Needless to say, it would be ideal to study an identical setting (a sequential tournament competition between two people who play a two-person game with a randomly determined order) performing a cognitive task rather than a noncognitive task. Luckily, this setting exists.

In a chess match, two players play an even number of chess games, typically either eight or ten games, against each other. One game is generally played each day, with one or two rest days during the duration of the match. The basic procedure establishes that the two players alternate the colors of the pieces with which they play. In the first game, one player plays with the white pieces and the other with the black pieces. In the second game, the colors are reversed, and so on. Who begins with what color is randomly determined, and this is the only procedural difference between the two players. According to the rules of FIDE (the Fédération Internationale des Echecs, the world governing body of chess), the order is decided randomly under the supervision of a referee. This random draw of colors, which is typically conducted publicly during the opening ceremony of the match, requires that the player who wins the draw plays the first game with the white pieces.

Hence, as in a penalty shoot-out, an explicit randomization method determines which player begins playing in a given role in a sequence of tasks or games where both players have exactly the same opportunities to play the same number of times in the same role, have the same stakes, and where all other circumstances are identical. As a result, as in a shoot-out, we should expect that two identical players have exactly the same probability of winning the match. That is, there is no rational reason why observed winning frequencies should be different from 50–50 in a large sample of chess matches. Yet Gonzalez-Díaz and Palacios-Huerta (2012) find that this is not the case. Instead, winning probabilities are about 60–40 in favor of the player who plays with the white pieces in the first and all the odd games of the match, and hence is more likely to be leading during the match.

*

The empirical evidence in this chapter shows that information on the performance of competing agents during the competition has an effect on noncognitive (soccer, basketball, weightlifting) and cognitive (chess) performance. Thus, as competitive situations that involve performing both cognitive and noncognitive tasks are ubiquitous in real life, the results may have broad applicability. There are, of course, numerous *strategic* reasons why in a sequential competition the order may give advantage to either a *first* mover or a *second* mover (see, e.g., Dixit and Pindyck, 1994; Cabral 2002, 2003). What the results in this chapter show is that there are, in addition, *psychological* reasons why leading or lagging may affect the performance of the competing agents.

An important consequence of these results is that a randomly determined order, which in a sequential tournament competition is obviously fair from an ex ante perspective, need not be ex post fair if it gives any type of advantage to a subset of the competitors.

The question then is this: How should the order of a sequential tournament competition between two agents be determined to make it both ex ante and ex post fair? Is it possible to improve upon the perfectly alternating order? A simple idea would be to change the type of tournament: Instead of sequential, make it simultaneous (e.g, the two teams in a penalty shoot-out may kick simultaneously in the two goals of the field). A similar alternative would be to keep it sequential but provide no information about the state of the competition until all the competitors have performed the same number of tasks. After all, it is *knowing* that one is leading or lagging that affects performance. For obvious reasons, either one of these alternatives is typically unfeasible or unattractive in sports, auctions, and other settings.

So, what can be done? Consider a sequential tournament where two players or teams A and B play against each other an even number of times. Say that a fair coin selects A to perform his or her task first and B second in the first two rounds. What should the order in the next two rounds be to attempt to make it ex post fair? Is there a way to improve the ex post fairness of the strict alternation of the order of play $A\,B\,A\,B\,A\,B\,A\,B\ldots$? Well, if the order $A\,B$ offers *any* kind of advantage to *either* player, then by reversing the order in the next two rounds, we tend to compensate that advantage. Doing so means that the resulting sequence in the first four rounds is $A\,B\,B\,A$. And, of course, this reversing is innocuous if no advantage existed in the first place. How about the next four rounds? The same principle applies: By reversing the order followed up to that point, we tend to compensate any potential advantage that might have been given to either one of the players until then. The resulting sequence is $A\,B\,B\,A\,B\,A\,A\,B$. And, again, reversing the order is innocuous if no advantage existed in the first place, that is, if $A\,B\,B\,A$ in the first four rounds already provides no ex post advantage to either player. Logically, we can apply the same principle ad infinitum and keep reversing the order followed from the beginning up to that point:

$$A\,B\,B\,A\,B\,A\,A\,B\,B\,A\,A\,B\,A\,B\,B\,A\ldots$$

This sequence is interesting, and it has a name: the Prouhet–Thue–Morse (PTM) sequence. Mathematician Axel Thue discovered it in Thue (1912) while studying avoidable patterns in binary sequences of symbols, e.g., 0 and 1. It is defined by forming the bitwise negation of the beginning:

$$\tau = 0\,1\,1\,0\,1\,0\,0\,1\,1\,0\,0\,1\,0\,1\,1\,0\ldots$$

where 1 is the bitwise negation of 0, 1 0 is the bitwise negation of 0 1, 1 0 0 1 is the bitwise negation of 0 1 1 0, and so on. Formally, the PTM sequence $\tau = (t_n)_{n \geq 0}$ is defined recursively by $t_0 = 0$ and $t_{2n} = t_n$, $t_{2n+1} = t'_n$ for all $n \geq 0$, where for $u \in \{0,1\}$ we define $u' = 1 - u$.

This sequence τ was already implicit in Eugène Prouhet (1851) and was later rediscovered by Marston Morse (1921) in connection with differential geometry. Worldwide interest in this sequence has developed during the past century as research has shown that it is ubiquitous in the scientific literature. In fact, this sequence occurs as the "natural" answer to various apparently unrelated questions, for instance, in combinatorics, in differential geometry, in number theory (e.g., the Prouhet–Tarry–Escott problem), in group theory (e.g., the Burnside problem), in real analysis (e.g., the Knopp function), in the physics literature on controlled disorder and quasicrystals, in music, in chess, in fractals and turtle graphics (e.g., the Koch snowflake), and in many other settings (see Allouche and Shallit (1999) for a survey).

Hence, the PTM sequence can also be the answer to an important problem in economics: How should the order of a sequential tournament competition between two agents be determined to make it both ex ante and ex post fair?

Unfortunately, the PTM ordering is not followed in tournament competitions, including major sports competitions, sequential auctions, and others. The closest we find is serving in tie-breaks in tennis where the order of serves one and two (*A B*) is reversed for serves three and four (*A B B A*), and then this sequence is repeated *A B B A A B B A A B B A* . . . until a player wins by a certain margin. The serving order in tennis would be perfectly fair ex post if any advantage given by the order in the first two serves, *A B*, is exactly compensated by having the order in the third and fourth serve reversed, *B A*. Of course, it is not known if this condition is empirically satisfied.

The PTM sequence, therefore, offers potential for improving the fairness of sequential tournament competitions.[9] It is important that the sequence has 2^{n+1} elements, $n \geq 0$, that is, that its first half is the negation of the second half. Otherwise, the full potential is not realized (e.g., in a soccer penalty shoot-out, the winner should be the best of $2^3 = 8$ penalty kicks or best of $2^4 = 16$, etc., not the best of 10 penalty kicks, as it currently is). Clearly, the margin of victory chosen to determine the winner is irrelevant for the ex post fairness of a sequence with 2^{n+1} elements.

9 Let $\Delta(\tau,n)$ denote the ex post difference in performance between the two identical subjects in a Prouhet–Thue–Morse sequence of 2^{n+1} elements, $n \geq 0$. Reversing tends to compensate any advantage if $|\Delta(\tau,n)|$ decreases with n. A necessary and sufficient condition for the PTM sequence to be ex post fair is that $\lim_{n\to\infty} \Delta(\tau,n) = 0$.

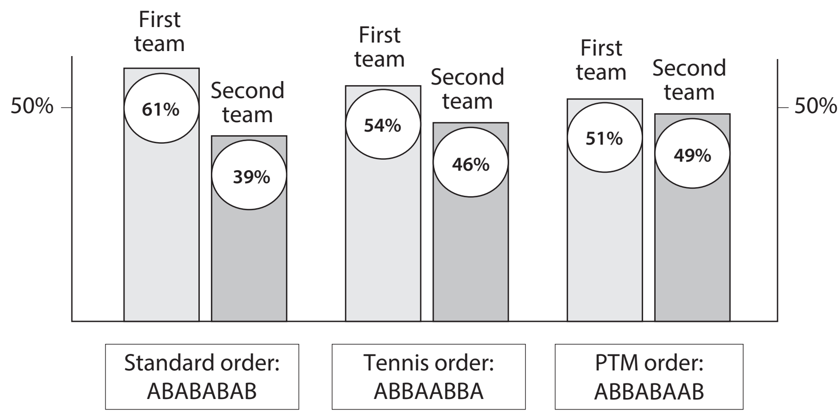

Figure 5.2. Winning frequencies using three different orders.

Since we are studying penalty shoot-outs in this chapter, it would be interesting to quantify the speed of convergence in this setting to produce an approximately fair outcome. How many rounds would be necessary to get "close enough" to 50–50? Is the number of rounds reasonable?

We explore this question with three experiments with professional players from Spain's La Liga (see figure 5.2). We implement penalty shoot-outs with three different kicking orders: In the first experiment, the sequence is *A B A B A B A B*; in the second, the order is the one followed in tennis: *A B B A A B B A*; and in the third, we follow the PTM sequence: *A B B A B A A B*. There are 200 shoot-outs in each experiment, each one involving 8 penalties, 4 per team, so that they can be perfectly compared. The order in each experiment is of course randomized.

The standard perfectly alternating order in the first experiment produces basically the expected advantage for the first kicking team: 61–39. Interestingly, when teams follow the tennis sequence, the advantage for the first kicking team decreases to 54–46, that is, from 22 percentage points it drops to just 8 percentage points. The advantage is further reduced if the PTM order is followed to just 2 percentage points: 51–49. Judging from these experiments, it appears that we do not need an excessive number of rounds to get reasonably close to 50–50, and so it seems quite feasible to improve the unfairness of the current perfectly alternating system in world soccer.

*

Sports competitions form an important class of fair division problems because sequences of strict alternation often give an unfair advantage to

Palacios-Huerta_1st-proof.indd 84 2/11/14 11:37 AM

one competitor. This chapter has shown that the advantage may be not only substantial but entirely psychological. There are other problems of fair division that also have the same structure and have already invoked the PTM sequence. Brams and Taylor (1999) invoked this sequence, but did not identify it as such, when allocating a contested pile of items between two parties who agree on the items' relative values. They suggest a method called *balanced alternation*, or *taking turns taking turns taking turns*, as a way to circumvent the favoritism inherent when one party chooses before the other. Levine and Stange (2012) proposed the PTM sequence as a way to reduce the advantage of moving first when sharing a meal (more precisely, in the Ethiopian Dinner game, in which two players take turns eating morsels from a common plate). Richman (2001) had already studied such equitable resource allocation problems, but he too did not identify the sequence as such at the time of publication. More recently, Cooper and Dutle (2013) show that two duelers with identical lousy skills (known as "Galois duelers" in honor of the famous mathematician Évariste Galois, who was killed in a duel at the age of 20) will choose to take turns firing according to the PTM if they greedily demand their chances to fire as soon as the other's a priori probability of winning exceeds their own.