Skill, Strategy, and Passion: an Empirical Analysis of Soccer*

Frédéric Palomino†

Luca Rigotti‡

Aldo Rustichini§

April 2000

Abstract

Sports provide a natural experiment on individual choices in games with high stakes. We study a game-theoretic model of a soccer match and then evaluate the ability of this model to explain actual behavior with data from 2885 matches among professional teams. In our model, the optimal strategy of a team depends on the current state of the game. When the game is tied, both teams attack. When losing, a team always attacks; when winning, it may attack early in the game, but starts defending as the end of the match nears.

We find that teams' skills, current score, and home field advantage are significant explanatory variables of the probability of scoring. We also find that when losing a team becomes relatively more likely to score. A team which is ahead, on the other hand, uses a conservative strategy very early in the match.

These results support the main conclusions of our model. They indicate that soccer teams behave consistently with rationality and equilibrium. However, we also find a strong and significant interaction between home field advantage and strategic behavior. Teams playing at home are more likely to score, unless they are already winning. This may be evidence that psychological factors are important in determining the game's outcome.

Keywords: Game theory, motivation, rationality, experiments, sports, soccer.
JEL codes: C73, C93, L83

*We thank Eric van Damme, Eddie Dekel, Patrick François, June K. Lee, Jan Magnus, Jean-François Mertens, and Chris Shannon for discussions.

†CentER, Tilburg University, and CEPR; `F.Palomino@kub.nl`.

‡CentER and Department of Econometrics, Tilburg University; `luca@kub.nl`.

§Boston University, and CentER, Tilburg University; `raldo@bu.edu`.

1 Introduction

We study a high stakes game between experienced players which occurs naturally: professional soccer. The aim is to test the ability to predict behavior using game theory and economic theory, much in the spirit of experimental economics. Our empirical test has three advantages over laboratory experiments. First, subjects are familiar with the game. Professional players and coaches have developed their knowledge of the game over many years. Second, subjects are highly motivated. Coaches are fired and players lose market value when performance is not acceptable. Third, a rich set of observations on strategic interaction is available. In a given year, more than 300 games are played in a top professional league.

Highly motivated professional athletes and wide availability of data make soccer a natural candidate for empirically testing predictions from a game-theoretic model in a natural environment. This paper is a first attempt in that direction.[1] Our analysis yields three main results. First, 'real world' data on strategic interaction of experienced players provide support for a game-theoretic explanation of behavior. The predictive power of rationality is good. Second, the surrounding environment has a strong influence on players, possibly consistent with psychological observations about behavior. Strategic behavior alone is not enough to explain what goes on. The third and most intriguing finding is that a consistent explanation of the data must allow the environment to interact with players' strategic choices. Strategic rationality and psychological elements are simultaneous and interacting forces in explaining behavior.

There are at least two lessons we draw from these result. The first is about method. Observing subjects in their natural environment, rather than in a laboratory setting, is possible

[1] In a spirit similar to ours, Walker and Wooders [1998] study the mini-max hypothesis in tennis games. Mini-max equilibrium play, refuted in several laboratory experiments (see O'Neill [1987] and Brown and Rosenthal [1990]), is instead supported by data from play of 10 tennis matches. Klaassen and Magnus [1998] also test several hypotheses concerning the distribution of points in tennis. Ferral and Smith [1999] test a model of championship series in football, basketball, and hockey. Unlike us, they focus on strategic choices *across* games rather than *within* a game.

and yields interesting results. In particular, the environment surrounding players choices may provide a substantial piece of the explanation for those choices. The second lesson is about theory. In the real world, behavior seems to depend on rational and psychological elements. These two aspects, though, cannot be studied separately since they clearly interact in determining a game's outcome. Understanding their interaction is the challenge for future theories.

What did we learn, and why is it important?

TV, radio and newspaper commentators of sports quite often mix detailed technical observations (on the choice of players, the layout of the team on the field, task assigned to different players) with observations of a psychological nature. These explanations have a common problem: they may give contradictory predictions. A team that has just allowed a goal is sometimes described as having a "reaction of pride" that makes it more likely to score than before. Other times, a team in exactly the same situation is described as "stunned", or "discouraged", and hence less likely to score.

In contrast, the game-theoretic model we develop gives a unique prediction: everything else being equal, a team that is down by a goal is, in equilibrium, *more* likely to score than a team in any other competitive position. The empirical analysis supports this prediction.

We analyze a game theoretic model of a soccer match, characterize its equilibria, and test its predictions with data from 2885 matches. Our focus is on teams' behavior at any given moment of the game; our measure of performance is the probability that a team scores a goal in that moment.

This model predicts that current score and time to the game's end influence teams' behavior. It also specifies how these strategic elements affect the probability of scoring during the game. We then test the predictions of the theory. The first main empirical result of the test is that the strategic elements are significant and important factors in explaining the probability of scoring. This is not, however, the end of the story.

Skill, strategy and passion: a quantitative estimate and a puzzle.

Theory and data detect three forces influencing performance: skill, strategy, and passion. *Skill* is a team's ability, the quality of players and coach. It is measured by long-run indices of attacking and defending technology like the number of goals made and allowed. *Strategy* is a team's choice to attack or defend in reaction to the game's score. It is measured by the relation a team's probability of scoring has with current score and time left until the game's end. *Passion* is the advantage a team has when playing with the support of the home fans. It is measured by the 'home field' advantage.

Table 1 illustrates in detail the impact skill, strategy, and passion have on the game. An entry under passion is the ratio between home and away probabilities of scoring; different entries are computed for each possible strategic environment (winning, losing, or tied). An entry under strategy is the ratio between probabilities of scoring corresponding to two strategic environments; different entries are computed for each possible state of passion (home or away). An entry under skill is the ratio between probabilities of scoring corresponding to two different ability values; they are computed for each state of passion.

Table 1: Determinants of the probability of scoring a goal in soccer games.[2]

		Home	Away		Losing	Tied	Winning		Home	Away
Strategy	$\frac{\text{losing}}{\text{winning}}$	2.20	1.39	Passion$_{\frac{\text{home}}{\text{away}}}$	1.59	2.02	1.00	Skill$_{\frac{\text{high}}{\text{low}}}$	2.32	2.21
	$\frac{\text{losing}}{\text{tied}}$	1.50	1.90							
	$\frac{\text{winning}}{\text{tied}}$	0.68	1.37							

Skill, strategy, and passion influence the probability of scoring a goal as follows:

i. skill differentials multiply this probability by a factor between 2.2 and 2.3;

ii. different strategic situations multiply it by a factor between 1.4 and 2.2;

iii. passion multiplies it by a factor between 1 and 2;

iv. strategy and passion interact in determining the probability of scoring since (a) the home-field impact varies with the current score, and (b) each current score impact changes when a team plays at home or away.

The quantitative effect of skill, the ex-ante most obvious explanation of performance, gives a benchmark to assess the importance of strategy and passion. Roughly, these three forces are equally important to understand behavior and performance.

Our results seem to vindicate the effectiveness of a pure game-theoretic analysis. They might justify the temptation of professional economists and game-theorists to dismiss explanations of psychological nature as unsubstantial. In this view, a team is like any player in any game, or any firm in any economy. And what is good to explain behavior and performance of a firm is good to explain behavior and performance of a soccer team.

This view, when combined with our results, considers passion an aspect of soccer's technology; that is, rather than the psychological effects provided by the home field advantage, there exist real, potentially quantifiable differences between home and away games. There are many ex-ante reasons to doubt this is the case. For example: noise aimed at hampering communications among players is unimportant since there is little 'play calling'; discomfort due to

[2]Losing (winning) indicates that the team is behind (ahead) by one or two goals in the score. The ratios for skill correspond to a team 40 percent better than the opponent versus one 40 percent worse; here, we only look at a tied game since different scores produce almost identical numbers. Full results and a detailed discussion are in section 4.

travelling is small since the cities where teams play are, in our data, relatively close; familiarity with the stadium is reduced since teams practice in a location different from the one where they play; the playing surface is natural grass for all games. But there are many other, more subtle, possible technological explanations of the home field advantage: sleeping in an uncomfortable hotel may make the visitors more tired, the referees may be influenced by local supporters, the dimensions of the field differ across stadiums, there are climatic differences among the cities where games are played, and so on. All these possibilities have one characteristic in common: they are constant throughout the game.

In our data, depending whether the team is winning, losing, or tying, the home field advantage varies. Roughly, Table 1 shows that playing at home doubles the probability of scoring when the game is tied, increases it by one and a half times when losing, and makes no difference when winning. The technological aspects of home field advantage are independent of the team's strategic situation, and thus cannot account for these differences.

Skill, strategy and passion: a possible interpretation.

Summarizing, our model and data show that skill and strategy, although very important, are not sufficient to give a full explanation of what is observed. If strategy and technology are not enough, we may be observing a psychological element of behavior. Commentators of sporting events often talk about a psychological 'extra man' effect the home fans may have on particular moments of a game. Spectators may influence players' behavior. Psychologists have documented a similar phenomenon: audiences modify behavior through 'social facilitation'.[3] When an audience is watching, the performance of very familiar actions improves. Little is known about social facilitation in strategic environments. Since the same audience may have

[3] Zajonc [1965] and [1969] are the seminal contributions on this topic. For a more recent survey, see Guerin [1993].

contrasting effects on players, the situation is necessarily more complicated and deserves further study.[4]

As economists and game theorists, we may be tempted to focus only on technology and strategy as a way to organize research by taking advantage of our professional competence. But measuring the importance of technological and rational versus psychological influences, we obtain similar orders of magnitude. Reason and emotion are empirically on equal grounds. Table 1 above also shows that they interact. The relative effect of different strategic situations changes when a team plays at home or away; the relative effect of playing at home or away varies with the current score.[5]

We conclude that psychology and rationality act as simultaneous determinants of the game's outcome. They interact in explaining behavior. For example, fans' support increases the probability of scoring when an additional goal is very important; that is, when the game is tied or the home team is losing. It has no effect when the home team is winning. These results, we think, point to the next challenge for economists and game theorists: build a theoretical model capable to explicitly take into account the interactions between emotion and reason these data highlight. Reversing the argument given above, one may now argue that what is necessary to explain behavior and performance of a soccer team is also necessary to explain behavior and performance of a firm. Therefore, a complete theory of the behavior of organizations, like a soccer team, cannot ignore any of them.

[4] In the case of zero sum games, for example, an audience friendly to one player is necessarily unfriendly to her opponent.

[5] Efficient division of labor may suggest to keep the study of reason and emotion separate if they do not interact. Only under such a strong separation assumption, the combined effect is determined by the algebraic sum of the parts. But if emotions influence behavior in a game 'additively', strategy must have the same effect wherever a team plays and passion must the same effect whatever the score. This additivity assumption is clearly falsified by the data.

Organization of the paper

The paper is divided as follows. Section 2 describes soccer, and models a match as a game. Section 3 characterizes the equilibria of this game. Section 4 contains the econometric results. Section 5 concludes. Proofs and extensions of the basic model are presented in section 6.

2 The Game of Soccer

In this section, we model soccer as a dynamic game between two teams. Each team chooses its strategy from a set of possible actions we call attacking intensities. They can be thought of as players' positions on the field as well as their mindset in playing the game; high attacking intensity means that a team focuses on offense more than defense. Strategies influence the probability that each team scores in any moment of the match. The state of the game is the current score and the time to the end. In any given moment, a team's strategy maps the state of the game into an attacking intensity. Equilibrium is, as usual, defined by optimal choices of strategies. We start with a brief description of the actual game.

ABC of soccer

The basic rules of a match are quite simple: eleven players on each side attempt to put the ball in the net of the opposing side; if they succeed, they score a 'goal'. The team with the highest number of goals at the end of the game wins; ties are possible.

Soccer is a low scoring sport. A single goal can change radically, and for a considerable amount of time, the strategic environment in which teams interact.

In domestic competitions, teams are rewarded with three points for a win, zero for a loss, and one point for a tie. Every match counts equally because national awards go to teams according to the sum of points collected on all a season's matches.[6] Therefore, we have repeated

[6] At the end of the year, points accumulated in each domestic league determine the national champion, participants in European tournaments in the following year, and teams relegated to a lower league.

observations of the same basic game.

Players' and coaches' incentives to perform well are strong since many soccer teams pay bonuses depending on match and/or season results. In addition, there is a very active and lucrative market for players and coaches, and the compensation depends strongly on past performance.

Players and strategies

There are two players, the teams, labelled $i = 1, 2$. The game is played in discrete time, with an instant denoted $t \in \{0, \ldots, T\}$, T being the final time. At each point in time, a team chooses the intensity by which it attacks from the strategy set $\{d, a\}$, where a is attack and d is defense. $\Delta(\{d, a\})$ is the set of mixed strategy. The choice of a pure strategy is perfectly observable by the other team. In keeping with the interpretation of "intensity of attack", we give the order $a \succ d$ to this set, and the natural (componentwise) partial order to the set of strategy profiles, denoted $s \equiv (s^1, s^2) \in \{d, a\}^2$. This order will be useful to discuss the technology of scoring.

The strategy set can be interpreted in two ways. It may measure the positioning of players on the field or it may indicate the mindset of players. Examples of the latter are deciding what to do when your team has the ball, or deciding how to react to ball possession of the opposing team. The assumption that the strategy set consists of two points is made for simplicity. In section (6.3), we consider the case of three actions.

Technology of scoring

Any team at any point in time can score a goal. Scoring is random, but the strategy choice of the teams affect the probability of a team scoring. These probabilities may be affected by factors other than strategies, like the home-field advantage or the skill of the two teams, which we introduce in the notation only at a later stage. Formally, for each pair of chosen strategies there is corresponding probability $p^i(s)$ that team i scores, and we write

$$p(s) \equiv (p^1(s), p^2(s)) \in [0,1]^2.$$

Because scoring is a rare event in soccer, both probabilities are typically very small (we will see that they are usually less than five per cent a minute).

We introduce now a natural assumption: increasing the intensity of attack of a team increases the probability of that team scoring, as well as the probability that the same team is scored against:

Assumption 2.1 (Monotonicity) *For $i = 1, 2$, p^i is increasing (that is, if $s \succeq s'$, then $p(s) \geq p(s')$).*

Note that this assumption defines an order over the quantities $p^1(s)$, except the relationship between $p^1(a,d)$ and $p^1(d,a)$, which is going to be determined later by the assumption 3.2. The different component parts of the monotonicity assumption seem all reasonable: for instance $p^1(a,d) > p^1(d,d)$ seems beyond doubt. The inequality $p^1(d,a) > p^1(d,d)$ seems reasonable, in light of the fact that team 2, by attacking, is making its own defense weaker. The third inequality, $p^1(a,a) > p^1(a,d)$ simply says that scoring against a defending team is harder than against an attacking team.

Payoffs

For each t, an integer n_t^i describes the total number of goals scored by team i until that instant. The game begins with a zero score, $(n_0^1, n_0^2) = (0,0)$. The payoff to each team depends on the final score according to the two functions G^1 and G^2 which depend only on the difference of the two scores:

$$(G^1(n_T^1 - n_T^2), G^2(n_T^1 - n_T^2)). \tag{2.1}$$

Let n_t denote the goal advantage of team 1: $n_t \equiv n_t^1 - n_t^2$. In this paper we consider the game as a zero-sum game, with payoffs

$$G^1(n) = 1, \text{ if } n > 0, G^1(n) = 0, \text{ if } n = 0, \text{ and } G^1(n) = -1, \text{ if } n < 0 \tag{2.2}$$

and $G(n) \equiv G^1(n) = -G^2(n)$. The case of payoffs $G^1(n) = 2$, if $n > 0$, $G^1(n) = 0$, if $n = 0$, and $G^1(n) = -1$, if $n < 0$, which corresponds more closely to the system used after the mid 90s, complicates the analysis but leaves the main result unchanged. For instance, since the payoff to victory is higher, the winning team will attack for a longer period.

The value function

A history at time t is the history of goals scored until that time. A strategy at time t is a function from the history in that period into the strategy set. For every period t and every pair (n^1, n^2) of goals scored until that time, there is a subgame beginning at t with that score, denoted by $\Gamma(t, n^1, n^2)$.

As usual, strategies may entail mixing over pure strategies. The mixed strategies of the two players in the subgame $\Gamma(t, n^1, n^2)$ are denoted $(\sigma^1(t, n), \sigma^2(t, n))$. Associated with $\Gamma(t, n^1, n^2)$, there is a value to the two teams of the score difference being n with only $T - t$ remaining to play. As the only determinant of payoffs is the final difference in the score this value only depends on n and t. We denote the value of the score being n at time t for team 1 and 2 respectively as $(v^1(t, n), v^2(t, n))$. Of course, $(v^1(T, n), v^2(T, n)) = (G^1(n), G^2(n))$.

The symmetric game

We begin with the case of teams of equal ability, i.e., which have identical probabilities of scoring. Formally we assume:

Assumption 2.2 (Symmetry) *$p^1(s^1, s^2) = p^2(s^2, s^1)$ for every (s^1, s^2).*

The case of non-symmetric games is considered in section 6.4. For convenience of exposition, we denote

$$\alpha \equiv p^1(a,d) = p^2(d,a); \delta \equiv p^1(d,a) = p^2(a,d); \tag{2.3}$$

$$A \equiv p^1(a,a) = p^2(a,a); D \equiv p^1(d,d) = p^2(d,d).$$

3 The equilibrium of the game

As the game is zero-sum, for any score difference n the value for team 1 of $\Gamma(t,n)$ is given by

$$v^1(t,n) \equiv \max_{\sigma^1} \min_{\sigma^2} E_{(\sigma^1,\sigma^2)} v^1(T,N); \tag{3.4}$$

the two strategies affect the probability distribution over a final difference in score N. The corresponding value for team 2 is $v^2(n,t)$. Since the game is zero-sum,

$$\text{for every } t \text{ and every } n, v(t,n) \equiv v^1(t,n) = -v^2(t,n). \tag{3.5}$$

Some properties of the value function are easy to derive, and useful to characterize the equilibria of the game. An immediate consequence of the fact that $v(T,\cdot)$ is increasing is that:

$$\text{for every } t, v(t,\cdot) \text{ is increasing.} \tag{3.6}$$

The equal skill assumption 2.2 has implications for the value function. Let π be the permutation $\pi(1) = 2, \pi(2) = 1$, and f the permutation on the strategy space:

$$f(\sigma) \equiv f(\sigma^1,\sigma^2) = (\sigma^2,\sigma^1). \tag{3.7}$$

$$p^i(s) = p^{\pi(i)}(f(s)), \text{ for every } i, s. \tag{3.8}$$

In addition, let $\hat{\sigma}(t,n) \equiv (\hat{\sigma}^1(t,n), \hat{\sigma}^2(t,n))$ denote the equilibrium of the subgame beginning at (t,n). Then from the symmetry assumption 2.2:

$$\hat{\sigma}(t,-n) = f(\hat{\sigma}(t,n)). \tag{3.9}$$

Denote by N the random variable taking values in the integers which describes the additional goals scored by the first team in the periods from t to T. It follows from (3.9) that

$$\begin{aligned} v(t,-n) &= E_{\hat\sigma(t,-n)}G(-n+N) \\ &= E_{f(\hat\sigma(t,n))}G(-n+N) \\ &= E_{\hat\sigma(t,n)}G(-n-N) \\ &= -E_{\hat\sigma(t,n)}G(n+N) \\ &= -v(t,n), \end{aligned} \tag{3.10}$$

and in conclusion

$$\text{for every } t \text{ and every } n, v(t,n) = -v(t,-n). \tag{3.11}$$

Note that in particular:

$$\text{for every } t, v(t,0) = 0. \tag{3.12}$$

From the equations defining the final payoff we get:

$$v(T,n) = 1 \text{ if } n > 0, v(T,n) = 0 \text{ if } n = 0, v(T,n) = -1 \text{ if } n < 0. \tag{3.13}$$

Finally the value satisfies the value function equation:

$$\text{for every } t \text{ and every } n, v(t-1,n) = \tag{3.14}$$

$$\max_{\sigma^1\in\Delta(\{d,a\})}\ \min_{\sigma^2\in\Delta(\{d,a\})}\ E_{(\sigma^1,\sigma^2)}[p^1v(t,n+1)+p^2v(t,n-1)+(1-p^1-p^2)v(t,n)].$$

Equilibrium with tied teams

To characterize equilibria, one needs to look at the possible score differences. We start from the case in which the game is tied.

Proposition 3.1 *Assume symmetry (assumption 2.2). If scoring when attacking while the other team defends is more likely than scoring when defending while the other team attacks, the equilibrium of a tied game has both team attacking; otherwise, they both defend. Formally,*

$$\text{if } \alpha > \delta, \text{ then } \hat\sigma(t,0) = (a,a); \text{ if } \alpha < \delta, \text{ then } \hat\sigma(t,0) = (d,d). \tag{3.15}$$

Proof. For every $(t, n) = (t, 0)$ the equilibrium solves the following.

$$\begin{aligned} v(t,0) &= \max_{\sigma^1} \min_{\sigma^2} p^1 v(t+1,1) + p^2 v(t+1,-1) \\ &= v(t+1,1) \max_{\sigma^1} \min_{\sigma^2} (p^1 - p^2) \end{aligned} \tag{3.16}$$

since $v(t+1, 1) > 0$. To solve the minimax problem in (3.16) one has to consider the matrix

	a	d
a	0	$p^1(a,d) - p^2(a,d)$
d	$p^1(d,a) - p^2(d,a)$	0

(3.17)

By the symmetry assumption 2.1, this matrix is equal to

	a	d
a	0	$\alpha - \delta$
d	$\delta - \alpha$	0

(3.18)

Now, one can easily conclude that $\hat{\sigma}(t, 0) = (a, a)$ if $P^1(a, d) = p^2(d, a) > p^2(a, d) = p^1(d, a)$, but $\hat{\sigma}(t, 0) = (d, d)$ if $p^1(a, d) = p^2(d, a) < p^2(a, d) = p^1(d, a)$. ∎

An equilibrium where neither team tries to win seems unreasonable, or at least uninteresting, particularly under the equal skill assumption. In the following we assume $\alpha > \delta$.

Assumption 3.2 (Attack is effective) *If symmetry (assumption 2.2) holds, then a team that attacks against a defending team is more likely to score than if it defends against an attacking team:*

$$p^1(a, d) > p^1(d, a).$$

Equilibrium with a winning team

We now consider the case in which the game is not tied.

Proposition 3.3 *If symmetry (assumption 2.2) and monotonicity (assumption 2.1) hold, then at equilibrium*

$$\hat{\sigma}(T-1,1) = (d,a) \tag{3.19}$$

Proof. Consider the period before the end with team 1 leading by one goal. We have:

$$\begin{aligned} v(T-1,1) &= \max_{\sigma^1} \min_{\sigma^2} p^1 v(T,2) + p^2 v(T,0) + (1-p^1-p^2) v(T,1) \\ &= \max_{\sigma^1} \min_{\sigma^2} (1-p^2) \\ &= 1 - \min_{\sigma^1} \max_{\sigma^2} p^2 \end{aligned} \tag{3.20}$$

If $A > \alpha$ and $\delta > D$ then defense is dominant for the first team, and the second attacks. ■

The switch between strategies

We have seen that in the late stages of the game a winning team chooses at equilibrium to defend. This does not mean, of course, that defending as soon as the team leads in the score is optimal. We prove now that in equilibrium a winning team continues to attack, until the end of the game is close enough. The time to switch depends on the score. This is the main implication of proposition 3.4.

Note first that

$$v(i, T-i+1) = 1 \text{ for every } i \leq T, \tag{3.21}$$

from the restriction that no team can score more than one goal in each period. Of course, any pair of strategies is an equilibrium at (i,n) if $n \geq T-i+1$, but we find it convenient to focus on

$$\hat{\sigma}(i,n) = (d,a), \text{ for every } n \geq T-i+1. \tag{3.22}$$

Now note that at $(t-1,n)$, with $n > 0$, the difference between the expected value of attacking and the expected value of defending against a team which is attacking is

$$(A-\alpha)v(t,n-1) + (A-\delta)v(t,n+1) - (2A-\alpha-\delta)v(t,n) > 0.$$

Let $\theta \equiv \frac{A-\alpha}{A-\alpha-\delta}$, and note that $\theta < 1/2$. It follows that attack gives a higher value if and only if:

$$\theta v(t, n-1) + (1-\theta) v(t, n+1) - v(t, n) > 0. \tag{3.23}$$

At the point $(T-i, i)$ the opposite of the inequality (3.23) is true, because of (3.21), so

$$\hat{\sigma}(i, T-i+1) = (d, a). \tag{3.24}$$

At $(T-2, 1)$ attack is better than defense if and only if

$$\alpha(A-\delta) - (A-\alpha)(1-\alpha) > 0,$$

but this term is positive at $\delta = 0, A = \alpha$, and is negative at $\delta = 0 = \alpha < A$. Therefore, for parameters values not excluded by our assumptions so far, both equilibria are possible in the interval between $(T-i, 1)$ and $(T-i, i-1)$. So an equilibrium where the winning team attacks after getting the lead, and then switches to defense is possible. We now show that after the winning team switches to defense, it keeps defending until the end or until the advantage is lost.

Denote by $C(t, n)$ the set of optimal strategies at (t, n) for the first team; $d \in C(t, n)$ means that defense is an optimal strategy at (t, n). We have the following lemma.

Proposition 3.4 *Suppose that for some t and some n,*

$$d \in C(t, n) \textit{ implies } d \in C(t', n) \textit{ for every } t' \geq t;$$

then

$$d \in C(t, n) \textit{ implies } d \in C(t, m) \textit{ for every } m \geq n.$$

The proof is in the appendix, section 6. There we also show how to compute a precise value for the time of switch in a continuous time version of the game. The value is reported in equation (6.36), and discussed in detail after (6.40).

A summary of the properties of the equilibrium

The equilibrium of this game has the following features. When the two teams are tied, the best response against an attacking team is attack. At this state of the score, the relevant probability is the difference between the probability of the two teams scoring. A team (say, the first) might choose to defend, but this choice would

i. reduce the probability of the other team scoring, since $p^2(d, a) < p^2(a, a)$,

ii. reduce the probability of the team itself scoring (this is the condition $p^1(d, a) < p^1(a, a)$),

The condition $p^1(d, a) < p^1(a, d) = p^2(d, a)$ implies that the second reduction more than compensates the first; hence, attack is a best reply to attack.

When the first team has one goal advantage the relevant probability is p^2, which the first team wants to minimize. This is clear in the last period: one additional goal of the winning team is useless. Hence, against an attacking team the best response is to defend, because this minimizes the probability $p^2(\cdot, a)$. The other team has to attack, because its objective is now to maximize the probability of scoring; an additional goal received does not make things much worse, and $p^2(d, d) < p^2(d, a)$. The key inequality in this argument is the assumption 3.2:

$$p^1(a, d) > p^1(d, a),$$

a possibly controversial statement among soccer fans, that we find reasonable. The opposite inequality might have been reasonable in the 60's, when defensive minded teams were very effective against attacking teams (Italian *contropiede*, or counterattack). To summarize, the main predictions of the equilibrium, if assumption 3.2 holds, are:

i. teams with equal score attack;

ii. when one of the teams has an advantage, the equilibrium has both teams attacking in the early stages of the game, and then the winning team defending and the other attacking in the later stages.

iii. a team which is losing is more likely to score, when the two teams are of equal ability;

4 Econometric Results

In this section, we present estimates of the probability of scoring a goal in any minute of a soccer match. This is a relatively novel idea. Most previous statistical studies on soccer have focused on the number of goals scored in a match.[7] The only similar example is Dixon and Robinson [1998], where a Poisson model is used to predict time of goals. As a matter of fact, most previous research on soccer has disregarded events during a game. The notable exception is Ridder, Cramer and Hopstaken [1994]; they study how the loss of a player affects a match's final outcome.[8] None of these papers attempts a connection between strategies and events during the game, therefore we are in a previously unexplored area.

Our objective is to test the main predictions of the model of the previous section. These are:

i. The probability of scoring is influenced by the current state of the game, measured by the difference between goals scored and the time left until the end.

ii. Losing teams adopt more aggressive strategies; losing has a positive influence on the probability of scoring.

[7] Among these, Reed, Pollard and Benjamin [1971], Maher [1982], Dixon and Coles [1997], and Palacios-Huertas [1998].

[8] When the referee shows someone a red card he is permanently expelled: his team plays the remainder of the match with one less player.

iii. Winning teams adopt more defensive strategies; winning has a negative influence of the probability of scoring.

As already noted, the model of the previous section makes predictions about optimal strategies. Under some assumptions about the linkage between probabilities and strategies, we derive the implications for the likelihood of observing a goal. Therefore, we also test restrictions on the scoring technology. The main of these is assumption 3.2, "Attack is effective"; it says the probability of scoring if a team attacks and the opponent defends is higher than the probability of scoring when it defends and the opponent attacks. Attack is effective is necessary to avoid the equilibrium of the symmetric game where teams defend if the score is tied. Although the low frequency of scoreless games in our sample, approximately 9%, seems incompatible with this equilibrium, a test of our model hinges on assumption (3.2) being confirmed by the data.

The main results of the empirical analysis are in section 4.4, where we describe in detail the probability of scoring a goal in a soccer game implied by our data. Here is a summary of the most important findings:

i. the probability of scoring depends on current score and time remaining;

ii. losing teams are more likely to score than winning teams;

iii. the magnitude of the previous effects depends on the home field advantage;

iv. teams' skills influence the probability of scoring;

v. country differences are sometimes relevant, but do not alter the previous picture;

These conclusions are consistent with our theoretical model and its underlying assumptions. On the other hand, size and direction of the home field advantage may be troublesome for a purely rational explanation of soccer teams behavior.

4.1 Match statistics

We use data on matches played in the Italian, English, and Spanish top professional leagues.[9] For each match, we know the names of the playing teams, their end of season statistics, and the minute in which they score a goal. The periods covered are 1995 to 1998 for Italy and England, and 1996 to 1998 for Spain. Overall, the data set includes 2885 matches

Table 2: scoring averages						
	Italy: 1044 games		England: 999 games		Spain: 842 games	
Goals per match	2.681		2.615		2.715	
Home		1.620		1.539		1.629
Away		1.061		1.076		1.086
Goals per minute	0.0296		0.0294		0.0305	
Home		0.0179		0.0173		0.0183
Away		0.0117		0.0121		0.0122

Table 2 presents some statistics on average number of goals per match and per minute. The home field advantage appears strong; home teams score about 0.5 more goals per game than away teams. Numbers are similar across countries.

The pattern of goals scored is illustrated by Figure 1.[10] It displays a well known regularity of soccer statistics: the frequency of goals scored increases with time, in an approximately linear fashion. Our theory and estimate provide a simple explanation of this regularity. As time goes by, the number of games that are tied decreases, and the sum of the probability of a goal being scored in a tied game is smaller than the total probability when one team is leading.[11]

[9] Match details were downloaded from internet sites of different news organizations. Although they are public, there is no unique source we could use to obtain them.

[10] Figure 1 displays a peak in the average number of goals scored at the end of each half. It happens because some extra time, called 'injury time', is added by the referee at the end of each 45 minutes half. For England and Spain, the data we obtained assign all goals scored in injury time to the 45th minute. For Italy, they register the actual time of the score, but not how long each match is played. To avoid measurement issues, we dropped the last minute of each half from the sample.

[11] This is made clear by the model with three strategy choices of section 6.3. There, the probability of observing a goal in a tied game, $2p^1(m,m)$, is smaller than the total probability when one team is leading, equal to $p^1(a,d) + p^1(d,a)$.

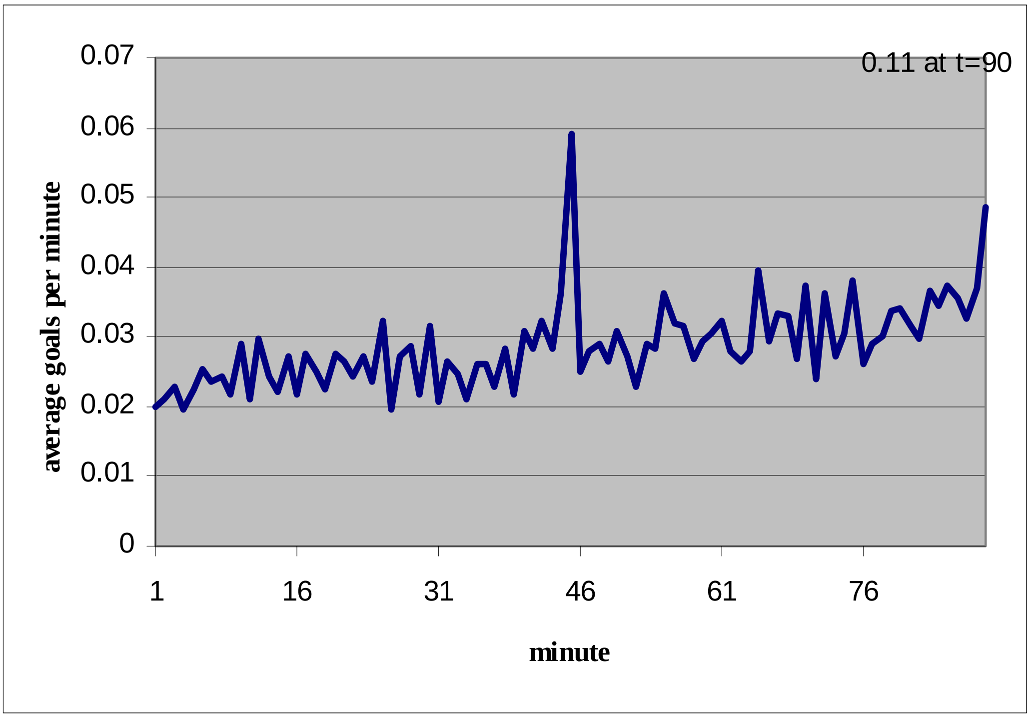

Figure 1: Frequency of goals in each minute of a soccer match.

Table 3: final score distribution				
	Number of Games		Percentage	
Tied	800		0.2773	
One goal	1046		0.3626	
Home		667		0.6377
Away		379		0.3623
Two goals	591		0.2049	
Home		372		0.6294
Away		219		0.3706
Three or more goals	448		0.1553	
Home		348		0.7768
Away		100		0.2232

Table 3 presents the most common final scores. The difference between goals scored and goals allowed is less than three in more than 80 percent of the matches. Differences of more than three goals are extremely rare.

4.2 Estimation Procedure

In this section, we describe our estimation of the probability of scoring. Let y_{itg} be a Bernoulli random variable equal to one if team i scores in minute t of match g and zero otherwise. Assume that the probability of observing a goal, denoted P_{itg}, is a function of a vector of explanatory variables x_{itg} and some unknown parameters β. In particular:

$$P\left[y_{itg}=1|x_{itg}\right]=E\left[y_{itg}|x_{itg}\right]=F\left(\beta x_{itg}\right)+\varepsilon_{itg} \tag{4.25}$$

where $F(\cdot)$ is the cumulative normal distribution function. If ε_{itg} is i.i.d. across teams, matches, and time, this is a standard probit random effects model. It can be estimated with maximum likelihood.

We make the following assumptions: team characteristics are constant during a season, and they are measured by overall performance in that season; match characteristics are represented by the opponent, and they are measured by opponent overall performance in the season and by a performance differential between team and opponent; time characteristics are given by current score and time left to be played.

We split the sample between teams playing home or away matches. Estimation over the entire sample may produce unreliable results. It assumes, unrealistically, independence among observations of the two teams playing a match.[12]

Regressors include a constant and 51 variables. These are divided in four groups: teams' skills, time elapsed, state of the match, and league (country) in which the match is played.

The skill variables are: *Offense*, measuring team i offensive strength, defined as goals scored in a season divided by matches played in that season;[13,14] *Defense*, measuring team i defensive weakness, defined as goals allowed in a season divided by matches played; *Opponent offense* and

[12]For completeness, we report estimation results over the entire sample in the appendix, section 6

[13]Season statistics are divided by games played because the number of teams differs across leagues.

[14]There is a mild endogeneity problem in using season averages because they include the goals of the current match. In a different version, we subtracted the outcome of current match without much effect on the estimates.

Opponent defense, measuring the opposition skills, defined as *Offense* and *Defense* for team i opponent; *Ability differential*, measuring how close the match is, defined as the absolute value of (*Offense-Defense-Opponent offense+Opponent defense)*.

The time variables are *Time*, a linear time trend, and *nn minutes left*, a dummy equal to one during the last *nn* minutes of the match and zero otherwise; *nn* is set equal to 75, 60, 45, 30, and 15. These dummies overlap, therefore the overall effect in the last 15 minutes is the sum of the five individual coefficients.

The state of the match is described by a set of dummy variables for the difference between goals scored and goals allowed. *Losing* equals one if team i is behind in the score by 1 or 2 goals at the beginning of minute t and zero otherwise. *Winning* equals one if team i is ahead in the score by 1 or 2 goals at the beginning of minute t and zero otherwise. These are interacted with all the time variables defined previously. *Losing by 3 goals* and *Winning by 3 goals* equal one if the score difference is not smaller than -3 or $+3$ respectively and zero otherwise.[15]

The country variables are *England* and *Spain*, and they equal one if the match is part of league in the corresponding country. These are also interacted with skills, time, and state of the match variables.

Given the set of dummies employed, the reference case is given by the first 15 minutes of a tied match played in Italy. Therefore, each dummy variable coefficient measures the marginal effect relative to this situation.

Finally, note that probit random effects estimation is appropriate only if the explanatory variables are not correlated with the error. Unobserved heterogeneity may thus be a problem.[16] Unobserved team, match, and time characteristics may affect the estimates' precision.[17] To

[15] In only 4 games, out of 2885, a lead of 3 goals did not end up in a win. Therefore, we think losing or winning by this margin implies the match is virtually over.

[16] One way to account for this is represented by conditional logit estimation. In our case, some independent variables (the current score ones) are a function of lagged values of the dependent variable (a team previous goals); hence, this method is not appropriate.

[17] Some examples are: ability variables beyond our performance measures, the particular conditions of the two

mitigate these effects, we compute robust standard errors and further allow clustering of observations for each team (in a season), for each match, or for each minute.[18] Therefore, we have four different standard errors for our estimate.

4.3 Estimation Results

First, we report a test that all the coefficients of the relevant current score variables (*Winning*, *Losing*, and their interactions with the time variables) are equal to zero. If this hypothesis is accepted, the data are not compatible with the theoretical model since the probability of scoring a goal does not depend on the current score. Table 4 below summarizes the results of the corresponding Chi-squared test. The null hypothesis is rejected; in all specifications, the probability the 18 variables are jointly equal to zero is smaller than 0.0015. Therefore, the probability of scoring in a soccer game does depend on the current score.

Table 4: joint significance of 18 current score variables.

	χ^2 tests with 18 degrees of freedom			
	Robust st. err.	Team clustering	Match clustering	Time clustering
Home matches	51.03	64.99	49.25	124.71
Away matches	43.89	53.12	41.32	52.52

Because the set of explanatory variables is large, an analysis of the results is somewhat cumbersome. We proceed as follows. First, we report estimates of the individual coefficients. Then, we describe in detail the probability of scoring a goal implied by these estimates. Results are divided in three blocks: skill and country variables first, time variables second, and current score variables last. We focus on marginal effects rather than coefficients and, to ease presentation, we only report p-values for robust standard errors with time clustering.[19] The full set of results is in the appendix.

teams on the day of the match, and players' fatigue.

[18]Additionally, we used a model where a common correlation structure across minutes of all matches is also estimated. Results are not substantially different, and are available upon request.

[19]Marginal effects are computed for 0 to 1 changes of the dummies. p-values refer to the underlying coefficient being zero.

Skill and country regressors.

Table 5 presents the estimates for skill and country variables. Only some of the skill regressors are significant. As expected, a team's offensive skills and its opponent defensive weakness influence positively the probability of scoring. The size of these effects is larger at home. Opponent offense and own defense do not matter. The ability differential has a significant impact only at home; more goals are scored in 'close' games in Italy and England, less in Spain. As for country effects, they are not significant in away matches, while at home teams are more like to score in England and Spain.

Table 5: regression results for skill and country variables.				
	Home matches		Away matches	
Variable	$\frac{\partial prob(goal)}{\partial Variable}$	p-value	$\frac{\partial prob(goal)}{\partial Variable}$	p-value
Offense	0.01346	0.000	0.00772	0.000
England	-0.00155	0.287	-0.00189	0.158
Spain	-0.00365	0.029	0.00007	0.947
Defense	0.00025	0.785	0.00015	0.824
Opponent offense	-0.00093	0.223	0.00018	0.810
Opponent defense	0.01242	0.000	0.00729	0.000
England	-0.00401	0.021	0.00075	0.598
Spain	-0.00477	0.004	-0.00079	0.649
Ability differential	-0.00226	0.001	-0.00055	0.399
England	0.00130	0.186	0.00031	0.768
Spain	0.00340	0.001	0.00098	0.338
England	0.00638	0.076	0.00352	0.249
Spain	0.00862	0.016	0.00031	0.914

Time regressors.

Time regressors have mixed result, as displayed in Table 6. Time elapsed does not influence the probability of scoring in away matches. At home, on the other hand, the linear time trend has a positive effect, partially mitigated by the time left in the game dummies. This trend is smallest in Italy and largest in Spain. These results indicate teams playing at home are more aggressive as time goes by.

Table 6: regression results for time variables.				
	Home matches		Away matches	
Variable	$\frac{\partial prob(goal)}{\partial Variable}$	p-value	$\frac{\partial prob(goal)}{\partial Variable}$	p-value
Time	0.00024	0.001	0.00007	0.253
England	0.00003	0.154	-0.00002	0.457
Spain	0.00005	0.031	0.00002	0.241
15 minutes left	-0.00338	0.048	-0.00157	0.299
30 minutes left	-0.00097	0.591	-0.00091	0.491
45 minutes left	-0.00534	0.001	0.00045	0.761
60 minutes left	-0.00395	0.012	-0.00095	0.448
75 minutes left	-0.00182	0.180	-0.00119	0.395

Score regressors.

Table 7 presents the results for the more relevant scoring variables. Losing influences positively the probability of scoring, even though the interaction with time tends to mitigate the overall impact. This is true both at home and away. Teams behind in the score seem to adopt more aggressive strategies exactly as our models suggests. These strategies become even more extreme toward the game's end. Winning influences negatively the probability of scoring. Interestingly, this effect appears smaller for away matches.[20] Both results do not appear to depend on the country the match is played in. Overall, current score has the predicted effect on strategies, but the impact is different for home and away teams.

For completeness, we report the estimate on score differences of 3 or more goals in Table 8. In most cases, they are not significant; the exception being losing by 3 goals which has a positive impact on the probability of scoring.

[20]This seems incompatible with a model where more aggressive play is chosen at home to entertain one's fans. That would imply the probability of scoring while winning should be larger at home than away.

Table 7: regression results for current score variables.				
	Home matches		Away matches	
Variable	$\frac{\partial prob(goal)}{\partial Variable}$	p-value	$\frac{\partial prob(goal)}{\partial Variable}$	p-value
Losing	0.01286	0.000	-0.00116	0.593
* Time	-0.00039	0.003	0.00031	0.004
15 minutes left	0.00896	0.002	-0.00207	0.337
30 minutes left	0.00448	0.148	-0.00437	0.019
45 minutes left	0.00484	0.112	-0.00398	0.036
60 minutes left	0.00957	0.007	-0.00333	0.116
75 minutes left	0.00156	0.674	-0.00396	0.093
England	-0.00178	0.258	-0.00143	0.173
Spain	0.00058	0.941	0.00076	0.828
Winning	-0.00829	0.013	-0.00635	0.000
* Time	0.00005	0.732	0.00031	0.006
15 minutes left	-0.00012	0.967	-0.00253	0.229
30 minutes left	-0.00232	0.300	-0.00238	0.263
45 minutes left	0.00379	0.167	-0.00358	0.118
60 minutes left	-0.00018	0.949	-0.00444	0.031
75 minutes left	0.00553	0.152	-0.00035	0.903
England	0.00015	0.923	0.00196	0.176
Spain	0.00031	0.866	0.00005	0.970

Table 8: regression results for score difference of 3 or more goals variables.				
	Home matches		Away matches	
Variable	$\frac{\partial prob(goal)}{\partial Variable}$	p-value	$\frac{\partial prob(goal)}{\partial Variable}$	p-value
Losing by 3 goals	0.02008	0.061	-0.00248	0.584
15 minutes left	-0.00079	0.914	0.00906	0.025
30 minutes left	-0.00589	0.460	-0.00202	0.558
45 minutes left	-0.00399	0.640	0.00428	0.492
England	-0.00527	0.372	0.00114	0.761
Spain	0.00058	0.941	0.00076	0.828
Winning by 3 goals	0.00066	0.900	-0.00596	0.370
15 minutes left	0.00321	0.420	0.00093	0.864
30 minutes left	-0.00639	0.027	0.00680	0.405
45 minutes left	0.00427	0.473	0.01024	0.458
England	0.00178	0.588	-0.00193	0.685
Spain	-0.00019	0.958	-0.00145	0.768

4.4 The Probability of Scoring in Soccer Matches

The estimation results enable us to recover the probability of scoring implied by the significant coefficients. We can then analyze how this probability changes with the state of the game, the home field advantage, or the teams' skills. Formally, the probability of scoring is given by the cumulative normal distribution

$$\int_{-\infty}^{\hat{\beta}x_{igt}} \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du,$$

where a dependent variable enters the formula only if the corresponding parameter estimate is significant at least at the 80 percent level. Dummy variables are included when the corresponding value is not zero. Table 9 presents the results of these calculations evaluating the skill variables at their means. Therefore, the numbers correspond to a game between teams of average skills.[21]

Table 9: probability of scoring a goal in each minute of an average skills game.

		Quarter of play						Game average
	Score	First	Second	Third	Fourth	Fifth	Sixth	
Home	Tied	0.0146	0.0147	0.0184	0.0166	0.0165	0.0185	0.0166
	Winning	0.0080	0.0080	0.0103	0.0116	0.0115	0.0180	0.0112
	Losing	0.0231	0.0254	0.0283	0.0237	0.0268	0.0212	0.0248
Away	Tied	0.0082	0.0082	0.0082	0.0082	0.0082	0.0082	0.0082
	Winning	0.0049	0.0078	0.0120	0.0126	0.0120	0.0182	0.0112
	Losing	0.0105	0.0161	0.0156	0.0159	0.0175	0.0182	0.0156

A goal is a rare event; the average probabilities of scoring are quite low (last column in Table 9). The highest, teams losing at home, is around 0.025. The lowest, teams tied away, is around 0.008. The average probability of scoring when losing is larger than the one when winning. This supports our attack is effective assumption, 3.2; scoring when attacking is more likely than scoring when defending. As time goes by, these numbers have an interesting pattern. Scoring is more likely late in the game. The increase over time is particularly significant when winning. Interestingly, winning teams switch to more defending strategies very early in the match. In the last quarter of play, their probability of scoring increases markedly. This could be due to reckless attacking by losing teams. As predicted by our theoretical model, at the end

[21]These numbers do not exactly correspond to a game between teams of equal (average) skills because the average skill differential, whose coefficient is negative and significant for home matches, is equal to 0.67. Setting this variable equal to zero would slightly increase the probability of scoring at home.

of the game a losing team must attack.

From Table 9, one can also deduce the size of home field advantage and the effect of changes in current score. To facilitate this task, Table 10 and Table 11 display the ratio between home and away probabilities and the ratio between probabilities corresponding to different scores. Passion and strategy have a large impact on the game's outcome. Unless a team is winning, playing at home increases the probability of scoring; it doubles in tied games, while it increases by more than 50 percent if a team is losing. Changes in the current score also imply large changes in the probability of scoring. These are more pronounced early in the game.

Table 10: effect of home field advantage.

	Quarter of play						Game average
Score	First	Second	Third	Fourth	Fifth	Sixth	
Tied	1.78	1.79	2.24	2.02	2.01	2.26	2.02
Winning	1.63	1.03	0.86	0.92	0.96	0.99	1.00
Losing	2.20	1.58	1.81	1.49	1.53	1.16	1.59

Table 11: effect of current score.

		Quarter of play						Game average
		First	Second	Third	Fourth	Fifth	Sixth	
Home	$\frac{\text{losing}}{\text{winning}}$	2.89	3.17	2.76	2.03	2.33	1.18	2.20
	$\frac{\text{losing}}{\text{tied}}$	1.58	1.73	1.54	1.42	1.63	1.15	1.49
	$\frac{\text{winning}}{\text{tied}}$	0.55	0.55	0.56	0.70	0.70	0.97	0.68
Away	$\frac{\text{losing}}{\text{winning}}$	2.15	2.08	1.30	1.26	1.46	1.00	1.39
	$\frac{\text{losing}}{\text{tied}}$	1.28	1.96	1.89	1.93	2.13	2.21	1.90
	$\frac{\text{winning}}{\text{tied}}$	0.60	0.94	1.46	1.54	1.46	2.21	1.37

Gauging the effect of the skill variables is less straightforward. One would like to assess how a global difference in skills between teams might affect the outcome of the match. To this end, we perform a comparative static exercise on all skill variables, taking average skills as reference point. First, we compute the probability of scoring if a team 20 percent better than average plays against a team 20 percent worse. Then, we compute the probability of scoring if a team 20 percent worse than average plays against a team 20 percent better. Finally, we compute the ratio between the first probability and the second; it measures the effect of being

40 percent better than the opposition versus being 40 percent worse.[22] The outcome of this procedure is displayed in Table 12. The effect of skills is large, since the probability of scoring more than doubles. It is not affected much by current score, and it slightly increases with home field advantage.

Table 12: effect of skills.

		Quarter of play						Game average
	Score	First	Second	Third	Fourth	Fifth	Sixth	
	Tied	2.35	2.35	2.29	2.32	2.32	2.29	2.32
Home	Winning	2.52	2.52	2.45	2.42	2.42	2.29	2.41
	Losing	2.22	2.20	2.17	2.22	2.18	2.25	2.20
	Tied	2.21	2.21	2.21	2.21	2.21	2.21	2.21
Away	Winning	2.32	2.22	2.13	2.12	2.13	2.04	2.12
	Losing	2.16	2.06	2.07	2.07	2.05	2.04	2.07

4.5 Summary of the empirical results.

The most interesting feature of tables 9 to 12 is that they enable a comparison among the effects of skill, strategy, and home field advantage. Overall, these three factors are simultaneous determinants of the behavior and performance in a soccer match. The magnitude of their contributions is roughly similar. The most intriguing message regards the interaction between home field advantage and strategic considerations.

Our theoretical model does not seem rejected by teams' actual behavior. They react to changes in the strategic environment along the lines predicted. More aggressive strategies are chosen when losing and more defensive ones when winning. The choice to defend when ahead seems to happen very early in the game. At home, the lowest scoring probability corresponds to being ahead in the score. Away, the lowest scoring probability corresponds to tied games. This suggests away teams are somewhat more conservative.

The effect of playing at home is very hard to reconcile with rational behavior. It influences

[22] In our sample, 1794 games (62%) have a skill differential at least this large.

positively the probability of scoring when losing or in a tied game. On one hand, there seem to be no 'technological' reasons for teams to do better on the home field. In our sample, geographical distances are relatively small; an away game involves very little travelling. Spectators do not directly influence the flow of play.[23] All matches are played on natural grass, and the field is of standard dimensions. On the other hand, the absence of a positive home field advantage when a team is winning rules out explanations based on keeping happy the home fans by playing more aggressively. In a sense, teams get extra aggressiveness from their supporters, but only if and when they need it to achieve success. That is, strategy and passion interact.

5 Conclusions

This paper has presented a game-theoretic model of a soccer match, characterized the equilibrium of the game, and tested it on a large sample of matches. The model predicts that teams' strategies in each moment, and therefore the probability of scoring a goal, depend on the score in that moment. It also describes how strategy and probabilities are affected by current score. For example, a losing team is more likely to score, and the probability that a winning team scores is smaller than when it is tied.

Overall, the model performs well in explaining observed data. Its qualitative predictions are confirmed (see for instance table 9 in the text). We can even explain a robust regularity of soccer statistics (the frequency of goals scored increases with the time elapsed in the match), which was so far a minor puzzle. The first conclusion, therefore, is that young and exuberant athletes are efficient professionals. They act according to economic reason when they play as well as when they bargain contracts with a team's owner or a sponsor.

A second conclusion, however, points out this is not the whole truth. The environment

[23] This is different from, say, American football and basketball, where many plays are 'called' by coaches, and the roar of the crowd can affect communications.

surrounding the game explains a large component of the data. Its effect appears consistent with emotional and psychological elements of behavior. Furthermore, these factors clearly interact with strategy and rationality.

This result has potentially wide implications. Soccer teams are examples of economic organizations who face each other in a very standardized, repeated, situation (a soccer match), which is therefore easy to study. Their behavior can provide insights on the way an economic organization works and in particular on the way strategic and emotional factors interact in its life.

In particular, we think there is a lot to be learned from a direct study of the interaction between passions and strategies, emotions and rationality. Since the three elements in the title of this paper coexist as important factors in explaining behavior, the challenge for future theoretical analyses is to explain their interaction.

6 Appendix

6.1 Proof of proposition 3.4.

The proof is by induction on t, going downwards from T to 0. The statement is clear for $t = T$. Suppose that $d \in C(t, n)$; by our assumption,

$$d \in C(t + 1, n), \tag{6.26}$$

and therefore by the induction hypothesis,

$$d \in C(t + 1, n + i), i = 1, 2. \tag{6.27}$$

Now equations (6.26) and (6.27) imply that

$$v(t + 2, n + i) \geq \theta v(t + 2, n - 1 + i) + (1 - \theta)v(t + 2, n + 1 + i), i = 0, 1, 2. \tag{6.28}$$

Also equations (6.26) and (6.27) imply that

$$v(t + 1, n + i) = E_d v(t + 2, n + i + \cdot)i = 0, 1, 2, \tag{6.29}$$

where E_d denotes the expectation with respect to the probability induced by the strategy d. But if we use the inequalities (6.28) we get:

$$E_d v(t + 2, n + 1 + \cdot) \geq \theta E_d v(t + 2, n + \cdot) + (1 - \theta)E_d v(t + 2, n + 2 + \cdot) \tag{6.30}$$

which, given the equalities (6.29), implies:

$$v(t + 1, n + 1) \geq \theta v(t + 1, n) + (1 - \theta)v(t + 1, n + 2). \tag{6.31}$$

This gives $d \in C(t, n + 1)$ as claimed. ∎

6.2 The Continuous-Time Game

In this section, we generalize the model described above to continuous time. This makes some of the arguments easier. Also, it allows a more precise study of the time of switching from attack to defense for the winning team, and some qualitative prediction on the way this time depends on parameters.

The game is defined as in section 2, but time t is the interval $[0, T]$. The process on the goals scored is now defined as a purely discontinuous Markov process. In each small time interval h, if the strategies are (s^1, s^2), the probability of scoring for team i is $p^i = p^i(s^1, s^2)$. So the score changes from n to $n+1$ with probability $p^1 h(1 - p^2 h)$, to $n-1$ with probability $p^2 h(1 - p^1 h)$, and remains unchanged with probability $p^1 p^2 h^2 + (1 - p^1 h)(1 - p^2 h)$. We denote by $\Delta_n v(t) \equiv v(t, n+1) - v(t, n)$, the right hand side "spatial" partial derivative.

Proposition 6.1 *The value function $(v(\cdot, n))_{n \in N}$ is the solution of the system of ordinary differential equations:*

$$-\frac{\partial v}{\partial t}(t, n) = \max_{\sigma^1} \min_{\sigma^2} E_{\sigma^1, \sigma^2}[p^1 \Delta_n v(t) - p^2 \Delta_{n-1} v(t)]; \tag{6.32}$$

$$v(t, 0) = 0, v(T, n) = 1, \text{ for all } n \geq 1. \tag{6.33}$$

The term in the maximization problem can be interpreted as the expectation of the inner product of the "spatial" derivative and the change in the state n.

Proof. Consider the discrete time problem, with time unit h, and write the functional equation for the value. Rearranging, dividing by h and taking limits yields the result. ■

An approximate game

A closed form solution of the continuous-time game is difficult. In this section we discuss briefly an approximate game, for which a closed form solution is possible. The game is defined as the original game: in particular, players and transition probabilities are the same. The only difference is that a team which, in any moment, reaches a two goals advantage wins the game.[24] It is clear that the value function for this game is the solution of the differential equation (6.32)

[24] In our sample of 2885 matches, there are only 96 cases in which a team leading by two goals ends up not winning at the end.

and the boundary conditions

$$v(t,0) = 0, v(t,2) = 1 \text{ for all } t, v(T,1) = 1. \tag{6.34}$$

This value function can now be explicitly determined. The solution for the interesting case (in which one of the two teams is leading by one goal) is presented in the following proposition.

Proposition 6.2 *In the game where the team with two goals advantage wins the value function is:*

$$\begin{aligned} v(t,1) &= \frac{\alpha}{\alpha+\delta} e^{(\alpha+\delta)(t-T)} + \frac{\delta}{\alpha+\delta}, \ \textit{if } T \geq t \geq t^* \\ v(t,1) &= (1/2-\theta)e^{2A(t-T)} + 1/2, \ \textit{if } t^* \geq t \geq 0. \end{aligned} \tag{6.35}$$

where

$$t^* = \max\{T - (\alpha+\delta)^{-1}[\log\alpha - \log((1-\theta)(\alpha+\delta)-\delta)], 0\}. \tag{6.36}$$

The equilibrium is $\hat{\sigma}(t,0) = (a,a)$, $\hat{\sigma}(t,2) = (d,a)$, *and*

$$\hat{\sigma}(t,1) = (a,a) \ \textit{if } t \leq t^*, \hat{\sigma}(t,1) = (d,a) \ \textit{if } t \geq t^*$$

and symmetrically for $n \leq 0$.

Proof. We first construct the value function, and then prove that it satisfies the differential equation (6.34). To construct the value function, consider separately the two differential equations

$$\frac{\partial v}{\partial t}(t,1) = (\alpha+\delta)v(t,1) - \delta \tag{6.37}$$

with boundary value $v(T,1) = 1$, and

$$\frac{\partial v}{\partial t}(t,1) = 2Av(t,1) - A \tag{6.38}$$

with boundary value $v(t^*,1) = v^*$ where v^* is a parameter determined in (6.39) below. The equality

$$-\alpha v(t^*,1) + \delta(1 - v(t^*,1)) = -Av(t^*,1) + A(1 - v(t^*,1))$$

determines the value at the switch point t^* between a and d; this equality is equivalent to:

$$v(t^*,1) = 1 - \theta. \tag{6.39}$$

This implies the value function has continuous first derivative in the time variable.

Paste the two solutions together, and obtain the value function as described in the statement of the proposition, and the value of t^*. Then one can immediately check that, since $2A > \alpha+\delta$, the value function we have defined satisfies the equation (6.34). ■

Let us discuss some comparative analysis of the solution which may be useful to understand the equilibrium. A key step in characterizing the equilibrium is the determination of the time where the winning team switches from attack to defense. This is done in the proof by solving separately two equations: one that gives the value of choosing defense, and the other the value of choosing attack (these are equations (6.37) and (6.38), respectively). Once this is done, one notes that the derivative of the solution of the first at $(T, 1)$ is α, and of the second is A; so the second value is, for t close to final time, smaller than the first. So for times close to the end clearly the winning team defends.

The winning team may attack in the initial times. However, as the solution for t^* indicates, there may be equilibria where the team with one goal advantage always defends. This may happen for instance if the values of α and δ are small.

The value of t^* is a good test for the predictive ability of the model. Note that

$$t^* = T + \frac{1}{\alpha+\delta}\log(1-\theta\frac{\alpha+\delta}{\alpha}). \tag{6.40}$$

Let us start considering what makes t^* large, i.e., what makes attack appealing. When A and α are close (so θ is small), attack is very appealing: in the extreme case, if $\theta = 0$, then $t^* = T$. This happens because there is very little gain in reducing the probability of getting a goal scored against. In fact when A is very close to α, then t^* is very high, irrespective of the absolute values of A, α, δ.

Now consider what makes t^* small, that is, what makes the winning team willing to defend early. Defense becomes appealing when:

i. α (which is the probability of getting a goal scored against while defending) is small relative to A, so there is gain from defending and keeping the advantage. Now θ is larger; recall that it cannot be larger than 1/2. But also

ii. the term $\frac{\alpha+\delta}{\alpha}$ (which lies between 1 and 2) is large, that is δ is close to α. Namely, the probability of scoring a goal while defending is not too low;

iii. the term $\alpha + \delta$ itself is not large.

To illustrate these results, consider the following numerical example. For $A = 3\%, \alpha = 2.5\%, \delta = 1\%$, so $\theta = 1/5$, the team which is winning attacks until 4 minutes from the end. This seems too close to the end. With $A = 2\alpha = 4\delta$, which gives a $\theta = 2/5$ and the log term approximately $-.4$ we get that the team defends $\frac{.4}{\alpha+\delta}$ minutes from the end. One needs a term $\alpha + \delta$ small to get a switch not too close to the end. For instance a value of $\alpha + \delta = 1\%$, which is probably too small, gives a switch to defense 40 minutes from the end. Although the order of magnitude is right, the model seems to predict that the winning team should switch to defense much later in the game than the evidence seems to indicate. This is in part consequence of the simplifying assumption that two goals give final victory, which makes attack more appealing. Another reason is that the strategy space consists only of two points.

In a more detailed version of this paper (Palomino, Rigotti and Rustichini (1998)) we present an extension that deals with the first problem. In the next section we present a model with larger set (three) of actions. In the extended version of the paper mentioned above, we also present a model with a continuum of actions.

6.3 Three actions

Suppose that each team can to choose an intensity of attack in the set

$$s \equiv (s^1, s^1) \in \{d, m, a\}^2.$$

where m now stands for "intermediate level of attack". This set is ordered in the natural way: $a \succ m \succ d$. We assume, as in the case of two strategies, that the probabilities are monotonic in this order, assumption 2.1. The analysis of the equilibrium is not significantly different. Consider first the case of the two teams at a tied score. The matrix 3.17 is now replaced by

	a	m	d
a	0	$p^1(a,m) - p^2(a,m)$	$p^1(a,d) - p^2(a,d)$
m	$p^1(m,a) - p^2(m,a)$	0	$p^1(m,d) - p^2(m,d)$
d	$p^1(d,a) - p^2(d,a)$	$p^1(d,m) - p^2(d,m)$	0

The equilibrium depends of course on the ordering of the elements in the matrix. If we assume

Assumption 6.3 (Attack is effective and risky) *A team that attacks against a defending team of equal skill is more likely to score than it is when defending against an attacking team; but the opposite holds when the other team is attacking with medium intensity. Formally:*

$$p^1(a,d) > p^1(d,a); \; p^1(a,m) < p^1(m,a); \; p^1(m,d) > p^1(d,m).$$

then

Proposition 6.4 *If symmetry (assumption 2.2), attack is effective and risky (assumption 6.3) and monotonicity (assumption 2.1) hold, the equilibrium of a tied game has both team attacking at the medium intensity m; formally,*

$$\hat{\sigma}(t, 0) = (m, m).$$

In the case of a team leading at a time close to the end, the equilibrium pair of strategies is again determined by

$$\min_{\sigma^1} \max_{\sigma^2} p^2.$$

The assumptions of proposition 6.4 imply that

$$\max\{p^1(m,d), p^1(d,d)\} < p^1(a,d) < \min\{p^1(a,m), p^1(a,a)\}$$

and therefore

Proposition 6.5 *If symmetry (assumption 2.2), attack is effective and risky (assumption 6.3) and monotonicity (assumption 2.1) hold, then at equilibrium*

$$\hat{\sigma}(T-1,1) = (d,a) \tag{6.41}$$

Our analysis of the data implies that the probability of scoring when tied is smaller than the probability of scoring when behind. These are the probabilities at equilibrium if

$$p^1(a,d) > p^1(m,m) > p^1(d,a). \tag{6.42}$$

6.4 The non-symmetric game

We consider the case where the probabilities of scoring a goal may be different, but we assume that the probability of scoring for the first team are higher for all strategy combinations. We denote

$$p_1(d,a) \equiv \delta_1, p_1(a,d) \equiv \alpha_1, p_2(d,a) \equiv \alpha_1, p_2(a,d) \equiv \delta_2,$$

$$p_i(d,d) \equiv D_i, p_i(a,a) \equiv A_i, \text{ for } i = 1,2, \tag{6.43}$$

so that in particular

$$A_i > \alpha_i > D_i, i = 1,2.$$

We make the following assumption:

Assumption 6.6 *There exists a $c > 0$ such that:*

$$\textit{for every } s, p_1(s) = p_2(\pi(s)) + c. \tag{6.44}$$

This means that the probability of the more skillful team is a parallel shift of each probability by the same magnitude. We denote v_c the value function when the probabilities satisfy (6.44).

Clearly, from the fact that $v(T,\cdot)$ is increasing,

$$v_c \geq v.$$

We denote by $\Gamma_c(t,n)$ the subgame with probabilities as in (6.44), at (t,n).

The equilibrium

First we prove that at $(i, T-i)$ the equilibrium is, as in the case of the Γ game,

$$\hat{\sigma}(i, T-i) = (d, a). \tag{6.45}$$

This follows from the fact the choice of a against a gives to the first player a value of 1 with probability $1 - A_2$, and a value of $v_c(i+1, T-i-1)$ with probability A_2, while the choice of d gives the same values with probability $1 - \alpha_2$ and α_2, respectively. Hence, the difference between the values from the first and the second choice is

$$(A_2 - \alpha_2)(v_c(i+1, T-i-1) - 1) < 0.$$

Also,

$$v_c(T-1, 0) = c, \hat{\sigma}(T-1, 0) = (a, a). \tag{6.46}$$

The argument is: compute the value to be

$$\max_{s^1} \min_{s^2} (p_1(s) - p_2(s)),$$

then use (6.44). Similarly,

$$v_c(T-1, 1) = 1 - \alpha_2, \hat{\sigma}(T-1, 1) = (d, a). \tag{6.47}$$

The argument is again an explicit computation and the equilibrium has already been found in (6.45).

6.5 Complete econometric results

In addition to the regressors listed in section 4, in the regression where all matches are considered we include *Home*, a dummy variable equal to one if the team plays home and zero otherwise; this is interacted with, current score, country, and time dummies.

Table A2: Regression results for home matches						
Number of obs	253880					
Log Likelihood	-21529.562	Pseudo R2	0.0141			
Observed probability of a goal	0.01697					
Predicted probability of a goal	0.01596					
Overall significance: chi2(51)			606.8	1376.9	605.5	2889.6
Variable	dF/dx	Estimate	Robust	Team	Match	Time
Offense	0.01346	0.3370	0.000	0.000	0.000	0.000
England	-0.00155	-0.0388	0.387	0.335	0.383	0.287
Spain	-0.00365	-0.0913	0.027	0.005	0.023	0.029
Defense	0.00025	0.0062	0.768	0.711	0.767	0.785
Opponent offense	-0.00093	-0.0233	0.238	0.215	0.230	0.223
Opponent defense	0.01242	0.3108	0.000	0.000	0.000	0.000
England	-0.00401	-0.1005	0.023	0.020	0.026	0.021
Spain	-0.00477	-0.1195	0.009	0.005	0.007	0.004
Ability differential	-0.00226	-0.0565	0.004	0.001	0.003	0.001
England	0.00130	0.0325	0.300	0.270	0.312	0.186
Spain	0.00340	0.0852	0.004	0.002	0.005	0.001
Losing by 3 goals	0.02008	0.3481	0.132	0.096	0.089	0.061
15 minutes left	-0.00079	-0.0201	0.913	0.908	0.913	0.914
30 minutes left	-0.00589	-0.1781	0.421	0.439	0.422	0.460
45 minutes left	-0.00399	-0.1124	0.645	0.623	0.659	0.640
England	-0.00527	-0.1554	0.380	0.361	0.363	0.372
Spain	0.00058	0.0142	0.937	0.937	0.935	0.941
Winning by 3 goals	0.00066	0.0164	0.899	0.897	0.894	0.900
15 minutes left	0.00321	0.0743	0.424	0.406	0.407	0.420
30 minutes left	-0.00639	-0.1954	0.074	0.056	0.071	0.027
45 minutes left	0.00427	0.0967	0.504	0.467	0.481	0.473
England	0.00178	0.0425	0.642	0.641	0.637	0.588
Spain	-0.00019	-0.0047	0.960	0.957	0.958	0.958
Losing	0.01286	0.2652	0.001	0.001	0.001	0.000
* Time	-0.00039	-0.0098	0.016	0.020	0.017	0.003
15 minutes left	0.00896	0.1862	0.023	0.025	0.022	0.002
30 minutes left	0.00448	0.1021	0.221	0.212	0.217	0.148
45 minutes left	0.00484	0.1104	0.205	0.212	0.214	0.112
60 minutes left	0.00957	0.2038	0.015	0.015	0.014	0.007
75 minutes left	0.00156	0.0379	0.690	0.678	0.686	0.674
England	-0.00178	-0.0465	0.273	0.296	0.281	0.258
Spain	0.00058	0.0142	0.937	0.937	0.935	0.941
Winning	-0.00829	-0.2292	0.003	0.002	0.004	0.013
* Time	0.00005	0.0011	0.736	0.714	0.741	0.732
15 minutes left	-0.00012	-0.0029	0.965	0.967	0.966	0.967
30 minutes left	-0.00232	-0.0610	0.382	0.372	0.394	0.300
45 minutes left	0.00379	0.0892	0.219	0.174	0.220	0.167
60 minutes left	-0.00018	-0.0044	0.950	0.944	0.950	0.949
75 minutes left	0.00553	0.1295	0.149	0.128	0.152	0.152
England	0.00015	0.0037	0.919	0.915	0.920	0.923
Spain	0.00031	0.0077	0.836	0.823	0.837	0.866
Time	0.00024	0.0061	0.003	0.004	0.003	0.001
England	0.00003	0.0008	0.190	0.173	0.199	0.154
Spain	0.00005	0.0013	0.036	0.025	0.031	0.031
15 minutes left	-0.00338	-0.0902	0.050	0.059	0.052	0.048
30 minutes left	-0.00097	-0.0245	0.593	0.594	0.596	0.591
45 minutes left	-0.00534	-0.1333	0.004	0.003	0.003	0.001
60 minutes left	-0.00395	-0.0956	0.019	0.016	0.018	0.012
75 minutes left	-0.00182	-0.0441	0.273	0.307	0.278	0.180
England	0.00638	0.1515	0.104	0.072	0.108	0.076
Spain	0.00862	0.1968	0.034	0.016	0.034	0.016
Constant		-3.0738	0.000	0.000	0.000	0.000

Table A3: Regression results for away matches						
Number of obs	253880					
Log Likelihood	-15611.301	Pseudo R2	0.0138			
Observed probability of a goal	0.01140					
Predicted probability of a goal	0.01069					
Overall significance: chi2(51)			456.6	627.6	460.3	1160.6
Variable	dF/dx	Estimate	Robust	Team	Match	Time
Offense	0.00772	0.2734	0.000	0.000	0.000	0.000
England	-0.00189	-0.0670	0.187	0.157	0.193	0.158
Spain	0.00007	0.0026	0.956	0.953	0.957	0.947
Defense	0.00015	0.0053	0.823	0.772	0.819	0.824
Opponent offense	0.00018	0.0062	0.784	0.809	0.791	0.810
Opponent defense	0.00729	0.2581	0.000	0.000	0.000	0.000
England	0.00075	0.0265	0.597	0.643	0.588	0.598
Spain	-0.00079	-0.0278	0.615	0.600	0.600	0.649
Ability differential	-0.00055	-0.0193	0.401	0.444	0.388	0.399
England	0.00031	0.0109	0.760	0.789	0.763	0.768
Spain	0.00098	0.0347	0.323	0.325	0.322	0.338
Losing by 3 goals	-0.00248	-0.0978	0.618	0.609	0.623	0.584
15 minutes left	0.00906	0.2432	0.033	0.037	0.034	0.025
30 minutes left	-0.00202	-0.0780	0.590	0.604	0.612	0.558
45 minutes left	0.00428	0.1312	0.538	0.526	0.543	0.492
England	0.00114	0.0385	0.740	0.735	0.753	0.761
Spain	0.00076	0.0260	0.825	0.835	0.836	0.828
Winning by 3 goals	-0.00596	-0.2918	0.399	0.379	0.382	0.370
15 minutes left	0.00093	0.0317	0.851	0.826	0.849	0.864
30 minutes left	0.00680	0.1930	0.404	0.409	0.399	0.405
45 minutes left	0.01024	0.2671	0.500	0.477	0.472	0.458
England	-0.00193	-0.0742	0.669	0.671	0.668	0.685
Spain	-0.00145	-0.0544	0.774	0.797	0.785	0.768
Losing	-0.00116	-0.0420	0.612	0.618	0.611	0.593
Losing * Time	0.00031	0.0111	0.005	0.006	0.006	0.004
15 minutes left	-0.00207	-0.0793	0.322	0.330	0.328	0.337
30 minutes left	-0.00437	-0.1803	0.028	0.032	0.030	0.019
45 minutes left	-0.00398	-0.1576	0.059	0.058	0.063	0.036
60 minutes left	-0.00333	-0.1273	0.125	0.100	0.132	0.116
75 minutes left	-0.00396	-0.1513	0.113	0.094	0.119	0.093
England	-0.00143	-0.0533	0.207	0.196	0.208	0.173
Spain	0.00076	0.0260	0.825	0.835	0.836	0.828
Winning	-0.00635	-0.2742	0.006	0.007	0.008	0.000
Winning * Time	0.00031	0.0109	0.016	0.011	0.016	0.006
15 minutes left	-0.00253	-0.0997	0.270	0.246	0.273	0.229
30 minutes left	-0.00238	-0.0921	0.320	0.280	0.313	0.263
45 minutes left	-0.00358	-0.1441	0.146	0.140	0.143	0.118
60 minutes left	-0.00444	-0.1827	0.059	0.061	0.060	0.031
75 minutes left	-0.00035	-0.0124	0.916	0.916	0.917	0.903
England	0.00196	0.0648	0.187	0.155	0.192	0.176
Spain	0.00005	0.0018	0.973	0.970	0.973	0.970
Time	0.00007	0.0025	0.301	0.303	0.309	0.253
England	-0.00002	-0.0005	0.447	0.430	0.458	0.457
Spain	0.00002	0.0008	0.295	0.247	0.294	0.241
15 minutes left	-0.00157	-0.0580	0.290	0.281	0.293	0.299
30 minutes left	-0.00091	-0.0326	0.551	0.531	0.548	0.491
45 minutes left	0.00045	0.0160	0.761	0.749	0.761	0.761
60 minutes left	-0.00095	-0.0331	0.492	0.493	0.497	0.448
75 minutes left	-0.00119	-0.0409	0.384	0.373	0.392	0.395
England	0.00352	0.1191	0.252	0.273	0.256	0.249
Spain	0.00031	0.0111	0.920	0.911	0.918	0.914
Constant		-3.1248	0.000	0.000	0.000	0.000

Table A1: Regression results for all matches						
Number of obs	507760					
Log Likelihood	-37165	Pseudo R^2	0.017			
Observed probability of a goal	0.01419					
Predicted probability of a goal	0.01314					
Overall significance: chi2(68)			1290.05	4581.32	1293.97	5454.59
Variable	dF/dx	Estimate	Robust	Team	Match	Time
Offense	0.00980	0.2901	0.000	0.000	0.000	0.000
Spain	-0.00173	-0.0513	0.882	0.896	0.885	0.896
England	-0.00173	-0.0511	0.232	0.253	0.223	0.226
home	0.00117	0.0346	0.408	0.412	0.408	0.421
Defense	0.00029	0.0087	0.100	0.014	0.102	0.077
home	-0.00017	-0.0051	0.128	0.007	0.129	0.062
Opponent offense	0.00011	0.0034	0.712	0.632	0.706	0.711
home	-0.00083	-0.0247	0.000	0.000	0.000	0.000
Opponent defense	0.00982	0.2906	0.872	0.869	0.872	0.880
Spain	-0.00278	-0.0824	0.894	0.896	0.895	0.886
England	-0.00154	-0.0456	0.020	0.019	0.014	0.018
home	-0.00014	-0.0041	0.168	0.161	0.167	0.160
Ability differential	-0.00113	-0.0333	0.516	0.518	0.525	0.450
Spain	0.00217	0.0642	0.041	0.037	0.039	0.033
England	0.00077	0.0228	0.986	0.986	0.986	0.986
home	-0.00041	-0.0122	0.662	0.686	0.668	0.627
Losing by 3 goals	0.00349	0.0934	0.018	0.013	0.018	0.034
15 minutes left	0.00698	0.1711	0.079	0.077	0.080	0.062
30 minutes left	-0.00385	-0.1314	0.960	0.963	0.961	0.963
45 minutes left	0.00095	0.0272	0.684	0.674	0.692	0.711
Spain	-0.00017	-0.0050	0.598	0.579	0.602	0.578
England	-0.00127	-0.0393	0.240	0.241	0.240	0.252
home	-0.00005	-0.0015	0.545	0.533	0.553	0.528
Winning by 3 goals	0.00043	0.0124	0.377	0.340	0.365	0.420
15 minutes left	0.00261	0.0714	0.287	0.245	0.283	0.244
30 minutes left	-0.00314	-0.1040	0.454	0.382	0.423	0.404
45 minutes left	0.00375	0.0997	0.003	0.003	0.003	0.001
Spain	-0.00058	-0.0174	0.030	0.036	0.029	0.017
England	0.00082	0.0237	0.422	0.410	0.424	0.406
home	-0.00134	-0.0414	0.045	0.042	0.041	0.042
Losing	0.00442	0.1215	0.073	0.062	0.073	0.053
* Time	0.00004	0.0012	0.273	0.283	0.295	0.282
15 minutes left	0.00151	0.0429	0.859	0.859	0.866	0.850
30 minutes left	-0.00179	-0.0557	0.448	0.434	0.448	0.373
45 minutes left	-0.00065	-0.0194	0.334	0.358	0.348	0.273
60 minutes left	0.00094	0.0273	0.745	0.744	0.748	0.727
75 minutes left	-0.00224	-0.0691	0.640	0.645	0.643	0.592
Spain	-0.00303	-0.0984	0.301	0.308	0.305	0.263
England	-0.00223	-0.0706	0.000	0.000	0.001	0.000
home	-0.00071	-0.0215	0.925	0.923	0.922	0.924
Spain	0.00252	0.0694	0.620	0.634	0.619	0.635
England	0.00120	0.0342	0.080	0.061	0.084	0.079
Winning	-0.00706	-0.2399	0.924	0.919	0.923	0.915
* Time	0.00016	0.0047	0.255	0.210	0.258	0.247
15 minutes left	-0.00121	-0.0373	0.835	0.823	0.832	0.815
30 minutes left	-0.00222	-0.0701	0.770	0.780	0.765	0.740
45 minutes left	-0.00005	-0.0015	0.826	0.821	0.829	0.822
60 minutes left	-0.00214	-0.0663	0.991	0.991	0.991	0.991
75 minutes left	0.00281	0.0791	0.456	0.422	0.460	0.445
Spain	0.00016	0.0048	0.487	0.486	0.494	0.520
England	0.00191	0.0539	0.207	0.162	0.212	0.143
home	-0.00031	-0.0092	0.980	0.978	0.980	0.978
Spain	0.00002	0.0007	0.241	0.238	0.244	0.189
England	-0.00138	-0.0427	0.267	0.240	0.273	0.259
Time	0.00016	0.0047	0.026	0.024	0.025	0.020
15 minutes left	-0.00243	-0.0762	0.163	0.208	0.166	0.120
30 minutes left	-0.00094	-0.0282	0.528	0.484	0.534	0.488
45 minutes left	-0.00234	-0.0692	0.606	0.597	0.613	0.582
60 minutes left	-0.00240	-0.0692	0.022	0.023	0.019	0.007
75 minutes left	-0.00148	-0.0426	0.046	0.034	0.047	0.046
Spain	0.00004	0.0011	0.457	0.454	0.450	0.453
England	0.00001	0.0002	0.205	0.224	0.209	0.218
home	-0.00001	-0.0003	0.026	0.006	0.028	0.020
Spain	0.00486	0.1348	0.069	0.033	0.065	0.028
England	0.00577	0.1610	0.000	0.000	0.000	0.000
Home	0.00712	0.2093	0.000	0.000	0.000	0.000
Spain	-0.00131	-0.0400	0.000	0.000	0.000	0.000
England	-0.00082	-0.0246	0.000	0.000	0.000	0.000
Constant		-3.2154	0.000	0.000	0.000	0.000

References

Brown J. and Rosenthal R., [1990]: "Testing the Minimax Hypothesis: A Reexamination of O'Neill's Experiment", *Econometrica*, 58: 1065-1081.

Dixon M. and Coles S., [1997]: "Modelling association football scores and inefficiencies in the football betting market", *Applied Statistics*, 46: 265-280.

Dixon M. and Robinson M., [1998]: "A birth process model for association football matches", *Journal of the Royal Statistical Society: Series D (The Statistician)* , 47: 523-538.

Ferral C. and Smith A.A., [1999]: "A sequential game model of sports championship series: theory and estimation", *Review of Economics and Statistics*, 81: 704-719.

Guerin B. [1993]: Social facilitation, *European Monographs in Social Psychology* Cambridge University Press, Cambridge, UK.

Klaasen F. J. C. M. and Magnus J. R., [1998]: "On the independence and identical distribution of points in tennis", CentER DP, Tilburg University.

Maher M., [1982]: "Modelling association football scores", *Statistica Neerlandica*, 36:109-118.

O'Neill B. [1987]: "Nonmetric tests of the Minimax Theory of two-person Zero-sum games", *Proceedings of the National Academy of Sciences*, 84: 2106-2109.

Palacios-Huerta I., [1998]: "One century of the world's most popular sport", *mimeo*, Stanford University.

Palomino F., Rigotti L., and Rustichini A., [1998]: "Skill, strategy and passion: an empirical analysis of soccer", CentER DP 98129, Tilburg University.

Reep C. and Benjamin B., [1971]: "Skill and chance in ball games", *Journal of the Royal Statistical Society (Series A)*, 131: 581-585.

Ridder G., Cramer J., and Hopstaken P., [1994]: "Down to ten: estimating the effect of a red card in soccer", *Journal of the American Statistical Association*, 89: 1124-1134.

Sahi S. and Shubik M., [1988]: "A model of a sudden-death field-goal football game as a sequential duel" *Mathematical Social Sciences*, 15: 205-15.

Szymanski S. and Smith R., [1997]: "The English football industry: profit, performance and industrial structure", *International Review of Applied Economics*, 11: 135-153.

Walker M. and Wooders J., [1998]: "Minimax Play at Wimbledon", *mimeo*, University of Arizona.

Zajonc R. B., [1965]: "Social Facilitation", *Science*, 149: 269-274.

Zajonc R. B., Heingartner A., and Herman E. M., [1969]: "Social Enhancement and Impairment of Performance in the Cockroach", *Journal of Personality and Social Psychology*, 13, 2: 83-92.