Does player specialization predict player actions? Evidence from penalty kicks at FIFA World Cup and UEFA Euro Cup

Does player specialization predict player actions? Evidence from penalty kicks at FIFA World Cup and UEFA Euro Cup

Luigi Buzzacchi* Stefano Pedrini†

November 2013

Abstract

Penalty-kicks are analysed in the literature as 'real life experiments' for assessing the use of rational mixed strategies by professional players. However, each penalty kick cannot be considered a repetition of the same event because of the varying background conditions, in particular the heterogeneous ability of different players. Consequently, aggregate statistics over datasets composed of a large number of penalty kicks mediate the behaviour of the players in *different* games, and the properties of optimal mixed strategies cannot be tested directly because of *aggregation bias*. In this paper we model the heterogeneous ability of players. We then test the hypothesis that differently talented players randomise over different actions. To this aim, we study a dataset that collects penalties kicked during shootout series in the last editions of FIFA World-Cup and UEFA Euro-Cup (1994-2012) where kickers are categorized as specialists and non-specialists. The results support our theoretical predictions.

JEL Classification: C72, C93, L83

Keywords: mixed strategies, penalty kicks, aggregation bias, players' specialization, action predictability

*Corresponding author: Politecnico di Torino, DIST, v.le Mattioli 39, 10125 Torino, Italy, (luigi.buzzacchi@polito.it) tel. 39 011 0907469 fax 39 02 700421226

†Università degli Studi di Bergamo, Department of Engineering, v.le Marconi 5, 24044 Dalmine (BG), Italy, (stefano.pedrini@unibg.it) tel. 39 035 2052010

1 Introduction

The individual behaviour of professional players during sports events has been often treated in the economic literature as a natural experiment for testing the empirical relevance of game theoretical predictions, particularly those of mixed strategy Nash equilibria in simultaneous two-persons, one-shot games. For this aim, sports events present a number of advantages: unlike the participants in many laboratory experiments, sport professionals are expert players, with evident incentives to behave rationally, and their actions are easily observable; moreover, unlike many alternative natural experiments, sports events are characterised by obvious payoff functions and simple interaction settings.
This strand of the economic literature is essentially based on the empirical analysis of datasets composed of large numbers of independent, similar events (especially penalty kicks) that are selected to obtain a suitable estimate of the behaviour and the payoff of the players. The outcomes are then compared in many ways with the theoretical predictions to assess whether the behaviour of expert players is consistent with the predictions of Nash equilibria.

Despite the advantages, the use of strategic interactions in sports events for assessing game theoretical predictions is not without difficulties. The discussion in the literature deals, in particular, with two critical issues: how a game model should represent observed situations (Is the game simultaneous or not? What are the strategic variables and the action sets of the players?) and how to interpret the empirical evidence. In the latter case, while game theoretical predictions hold for individual observations, empirical tests can only be performed on samples of *heterogeneous* observations. Given that different players are likely to be characterised by different payoffs or even different strategy sets, it's easy to understand that the properties of aggregated estimates do not necessarily coincide with the properties required for individual behaviour.

Restricting our attention to penalty kicks in soccer, some of the questions addressed above find adequate answers in the literature. The evidence that professional players actually behave at the equilibrium as predicted by theory is quite extensive. Moreover, the methodologies proposed so far for testing simultaneity and serial independence of actions are widely accepted and penalty kicks are normally recognised as simultaneous one-shot games. Less attention has been paid in this literature to the analysis of the effects of aggregation of heterogeneous events, in particular considering heterogeneity determined by differences in players' quality. The scarce attention to the issue is justified by the fact that the samples analysed so far in the literature present a population that is not very heterogeneous, as will be discussed below.

In this paper we provide an in-depth analysis of aggregate behavioural outcomes when players are very heterogeneous. For this purpose, we analyse a set of penalty kicks selected to obtain a sample of heterogeneous players acting in a relatively uniform environment to highlight the role of the different ability of the players. Our dataset is composed of every penalty kicked during shootout series at FIFA World Cup and UEFA Euro Cup

from 1994 to 2012.[1] Shootout series produce valuable data because players other than "penalty specialists" are called to kick while guaranteeing more uniform incentives and fatigue for every player.[2]
Our main result is that when the ability of the kickers is sufficiently heterogeneous, modeling the actions of the players with only two or three actions associated to the side of the kick - as usually happens - is not sufficient to represent the strategic space. We show that differently skilled players not only randomise with different probability distributions over the same actions, but they also randomise over different actions. This evidence can't be obtained within the framework used in the previous literature.

In the paper's second section, we survey the economic literature on penalty kicks. In the third section, we present our dataset. In the fourth section, we present a series of theoretical predictions and discuss the results. Section five briefly concludes.

2 The economic literature on penalty kicks

The empirical significance of theoretical predictions concerning mixed strategy equilibria has been often analysed in the economic literature addressing professional sports events, and in particular penalty kicks in soccer matches. Other sports applications include, among others, first serves in tennis (Walker and Wooders [2001] and Hsu et al. [2007]), soccer shots during live action (Moschini [2004]), pitches in baseball and play choice in football (Kovash and Levitt [2007]).

A penalty kick can be simply treated as a generalised matching pennies (**GMP**) game. A GMP is a constant-sum, one-shot game. The players simultaneously choose an action out of the two available (for both players) and if the choices "match", one of the players obtains a payoff higher than the payoff obtained if the choices do not "match". The opposite holds for the rival. Obviously, a GMP presents no Nash equilibria in pure strategies.
In a penalty kick game, the players are the kicker (**K**) and the goalkeeper (**G**), and their choices are deciding where to kick and where to jump, respectively. The game varies from a GMP for the following reasons:

- there are actually more than two available actions: K can direct the shot toward any point in the goal area, adjust the power of the kick, etc.; G can direct his jump

[1]A penalty shootout is a method used to decide which team progresses to the next stage of a tournament (or wins the tournament) following a tied game. In a nutshell, the shootout series is composed of five penalties for every team, but every single penalty must be kicked by a different player, so that even non-penalty-specialists are involved. The team that scores more penalties obviously wins. Details concerning specific rules and the history of shootouts can be found at `en.wikipedia.org/wiki/Penalty_shoot-out_(association_football)`

[2]Shootout penalties are kicked at the end of the match and every single kick is obviously crucial. On the contrary, with the relevant exception of the paper by Apesteguia and Palacios-Huerta [2010] discussed in sec. 3.1, the datasets used in the literature refer to penalties kicked during national league matches. These penalties are usually kicked by the team-specialist and can be assigned at any moment of the match. The outcome of these penalties can be more or less crucial, depending on the score and the time remaining to the end of the match.

in many different directions, decide where to stand on the goal line, etc.;

- the action choice of each player does not necessarily result in the outcome observed, because the actual action of the player could, more or less, fail;[3]
- every action profile does not determine a certain payoff but only a scoring probability. For example, when the actions of the players match (i.e. they choose the same region of the goal area), the goalkeeper may not be able to clear the ball; alternatively, when the actions do not match, the ball could miss the goal. Intuition (and the empirical evidence) suggests that the probability of scoring a goal increases, ceteris paribus, when the actions of the two players do not match (and decreases when they do match).

As for the first point, the seminal paper of Chiappori et al. [2002] first represents the action set of each player as composed of three actions: the kicker and the goalkeeper can kick/jump to their left, central or right side.[4] In most of the works that followed Chiappori et al., the pure strategy set has been further reduced to two elements: left and right.[5] To be more precise, the composition of the action set of the players usually represented in the literature takes into account the fact that the game is not symmetric because the kicker has a preferred or '*natural*' side. According to conventional wisdom, when kicking to the converse side of his strong foot, a kicker obtains a higher scoring probability because, ceteris paribus, shots are placed with more strength and/or precision; this fact is explained by the biomechanical characteristics of the movement. Given this asymmetric performance, the kicker must actually decide wether to kick on his '(N)atural' side (left for a right-footed kicker and right for a left left-footed kicker) or on his '(O)pposite' side.

The reduced form of the game (2×2 actions) is then represented in tab. 1. Since the payoffs for K and G sum to one, only the payoff for K is reported. The ranking of the parameters is $1 > \pi_N > \pi_O > P_N > P_O > 0$ (following empirical evidence, the advantage of wrong-footing the goalkeeper is assumed to be larger than the natural side advantage). The mixed strategy equilibrium of the game can be easily obtained.

[3] A kick out of the goal is an example of such a failure, even if often reveals at least the broad intended shot direction. The same is less easy to detect for centre kicks, which could be "failed" left or right kicks.

[4] Chiappori et al. and the following literature therefore choose to consider only the 'horizontal direction' of the kick (or jump), deliberately ignoring other decisions, such as the vertical direction of the ball or the power of the kick. The only empirical analysis taking into account the vertical direction of a penalty kick is provided by Bar-Eli and Azar [2009]. This choice can be easily understood considering that during data collection (data are usually obtained by direct inspection of the videos of soccer matches), it is difficult to identify vertical directions and almost impossible to measure the power of the kick.

[5] Palacios-Huerta [2003, footnote 11] decides to drop the pure strategy of centre kicks/jumps because of the small proportion observed in his empirical analysis. The same applies in Baumann et al. [2011]. Leininger and Ockenfels [2007], on the contrary, assign a relevant strategic role to central kicks. They observe an increase of the scoring probability in the German Football League (Bundesliga) in 1963-1990, and they attribute this phenomenon to a behavioural innovation that took place when kickers, during the '70s, recognised the strategic value of occasionally kicking to the centre. On the empirical side, the decision to limit the action set of the players to just two pure strategies eliminates any ambiguity that could be introduced when defining the width of the area of the goal to be considered 'central'.

		Goalkeeper (G)	
		Natural	Opposite
Kicker (K)	Natural	P_N	π_N
	Opposite	π_O	P_O

Table 1: Asymmetric 2×2 penalty kick game

Starting from the paper of Chiappori et al. [2002], the typical research strategy in the literature - aimed at assessing the equilibrium behaviour of professional players - selects a large number of penalty kicks, observes the action of the players, measures the frequency of the different actions and the scoring probabilities for every action-profile, and then controls whether the estimated figures are coherent with the theoretical predictions. To achieve this aim, the natural test to perform is based on the 'Fundamental Lemma' of mixed strategies. This lemma presumes that, at the equilibrium, the expected payoffs of every pure strategy that is randomised with positive probability must be equal.[6]

Even in the simplified model represented in tab. 1 the actual consistency of the observed behaviour of professional soccer players with the predictions of mixed strategy Nash equilibria is not easy, because of observations heterogeneity. If the datasets were all composed of homogeneous observations, the obvious strategy for verifying that players actually play Nash equilibrium strategies would be to check whether the expected payoff for both the kicker and the goalkeeper that choose to kick/jump natural is not significantly different from the expected payoff they obtain when choosing to kick/jump opposite (applying the Fundamental Lemma). However, every single penalty kick observed is indeed a *different* game because P and π depend on the ability of the two players and on several other variables (such as the athletic/psychological conditions of the players or field circumstances[7]). If a reasonable number of observations is needed, one has to accept a certain level of heterogeneity. The average scoring probability, for example, is then only a mean of the different values assumed by P and π in every individual observation. As a consequence, because different P's and π's characterise every observation, the average scoring probabilities do not need to comply with the Fundamental Lemma even at equilibrium. For the same reason, average payoffs also do not need to comply with the ranking of the payoffs expected for every single observation.[8] If the observations are heterogeneous, an *aggrega-*

[6] As said, the literature usually assumes that no one player is able to anticipate the other by perceiving some clue about his choice, so that the game is indeed simultaneous; moreover, it is assumed that the game is one-shot, in that the memory of past repetitions does not induce dynamic strategies. The empirical evidence confirms these assumptions (Miller [1998], Chiappori et al. [2002] and Palacios-Huerta [2003]).

[7] Apesteguia and Palacios-Huerta [2010] analyse a large set of penalty shootouts in order to understand the effect of psychological pressure on the performance of the players. Their main findings are discussed in sec. 3.1. Dohmen [2008] documents that penalty kickers are more likely to choke on a penalty kick when the match takes place in their home stadium.

[8] To be more precise, it is easy to prove that the ranking of average payoffs respect the expected ranking for every single observation if P's and π's differ only by a constant among players. The same applies for the Fundamental Lemma.

tion bias may arise.
Assume, for example, that the population of the kickers is characterised by two types of players, differing only by the value of P_N. Those characterised by higher values of P_N kick more often on their natural side and goalkeepers facing such players jump more often on the natural side. As a consequence, among the shots on the natural side, those kicked by high-P_N players are overrepresented; consequently, the choice to kick on the natural side seems to present a higher (average) scoring probability even at the equilibrium, thus apparently violating the Fundamental Lemma.

To avoid a direct test of the Fundamental Lemma, Chiappori et al. [2002, Prop. 3] find a number of properties of the equilibrium that do not depend on the composition of the population, but only require that the payoffs *for every player* follow the ranking indicated above ($1 > \pi_N > \pi_O > P_N > P_O > 0$).[9] For example, at the equilibrium, kickers must kick opposite less frequently than goalkeepers dive opposite. All of these properties are confirmed by the empirical evidence, which is provided by a dataset composed of 459 penalties kicked in the Italian and French first leagues in 1997-2000.
Coloma [2007] confirms on the same dataset the results obtained by Chiappori et al., proposing an original methodology that uses a simultaneous regression approach.
Palacios-Huerta [2003] collects a dataset composed of 1417 penalties kicked in different leagues (most Spanish, Italian or English) in 1995-2000. This database is so large that it is possible to test whether the scoring probabilities are identical across pure strategies for *individual* players: 42 players (both kickers and goalkeepers) were involved at least in 30 penalty games. Palacios-Huerta finds that the null hypothesis of equal scoring probabilities cannot be rejected for 39 players at the 5% significance level.

In summary, the literature confirms that soccer professionals' behaviour is remarkably consistent with equilibrium choices.[10] These results are robust to players' heterogeneity. No evidence that different players behave differently is provided in the literature. The only exception is Baumann et al. [2011], who propose a probit analysis on a dataset of 999 penalties kicked in the German Bundesliga in 1995-2007. They observe that high quality players choose more frequently the natural side, a choice that makes them more predictable; however, high quality players also present a higher scoring probability. The result points out a bias towards the natural side that increases with the ability of the player.
None of the authors cited challenge the implicit assumption in the literature that different players randomise over the *same* action set.[11] It is important to remember again here

[9] Actually, another assumption is needed: $\pi_N - P_N$ must be lower than $\pi_O - P_O$ for every player. More properties characterising the mixed-strategy Nash equilibrium can be derived under the hypothesis that heterogeneity regards only one of the two players.

[10] Bar-Eli et al. [2007], however, discuss possible non-optimal behaviour of goalkeepers driven by 'inaction aversion'.

[11] Chiappori et al., (Prop.1) find that a "restricted randomisation" is possible when the scoring probability of kicking central is sufficiently low. In this case, the kicker never kicks to the centre and the goalkeeper never remains at the centre. The authors do not provide any empirical test about this prediction because of the small number of central kicks in their dataset.

that the heterogeneity of the players in the samples analysed in the cited literature is likely to be fairly limited. In our paper, we will discuss this assumption and provide empirical evidence derived from a dataset where heterogeneity across players is much more pronounced.
Similarly, the literature does not discuss the fact that the direction of the kick that is observed could be an imperfect realisation of the intentions of the kicker. In other words, every dataset always contains a number of 'observational mistakes'. A corollary to the previous observation is that, not surprisingly, in most of the literature on penalty kicks, a blocked kick is treated the same way as an outside-the-goal kick. The following empirical analysis will also distinguish the information associated to these two types of events.

3 Data and variables

Our dataset comprises all the penalties kicked during all shootout series at the FIFA World Cup (from 1994 to 2010) and the UEFA Euro Cup (from 1996 to 2012). Descriptive statistics of the dataset are provided in tab. 2.[12]

Tournament	Playoff matches	Matches ended with shootout	Penalties (a+b+c)	Goals (a)	%	Kicks saved by the goalkeeper (b)	%	Kicks out of the goal (c)	%
FIFA 1994	16	3	29	18	62.1%	7	24.1%	4	13.8%
UEFA 1996	7	4	42	37	88.1%	4	9.5%	1	2.4%
FIFA 1998	16	3	28	20	71.4%	7	25.0%	1	3.6%
UEFA 2000	7	1	8	4	50.0%	3	37.5%	1	12.5%
FIFA 2002	16	2	19	13	68.4%	3	15.8%	3	15.8%
UEFA 2004	7	2	26	20	76.9%	2	7.7%	4	15.4%
FIFA 2006	16	4	33	21	63.6%	8	24.2%	4	12.1%
UEFA 2008	7	2	16	10	62.5%	4	25.0%	2	12.5%
FIFA 2010	16	2	18	14	77.8%	2	11.1%	2	11.1%
UEFA 2012	7	2	18	12	66.7%	3	16.7%	3	16.7%
Total	115	25	237	169	71.3%	43	18.1%	25	10.6%

Table 2: Descriptive statistics of the dataset

Each penalty is an observation. For each observation, we have collected data on the identity of the kicker and the goalkeeper, as well as the foot used by the kicker for the shot and the action chosen by both players. For the action sets, we have made the following choices:

[12]The shootout series take place only at the end of playoff matches that end in a tie. It must be remarked that there need not to be ten penalties in each series. Teams take turns to kick until each has taken five kicks. However, if one side scores more goals than the other could possibly reach with all of their remaining kicks, the shootout ends regardless of the number of kicks remaining. On the other hand, if the five penalty series end in a tie, further penalties are kicked until one team scores and the other team does not.

- We have vertically divided the goal area into two equal parts. Kicks have been classified depending on the half of the goal area where they have been directed.[13] Combining this information with the foot used for the kick allows us to classify actions as natural side shots (N) or opposite side shots (O);

- For the goalkeeper, three possible alternatives are evaluated: the goalkeeper jumps to his left, the goalkeeper jumps to his right or the goalkeeper stands still. Similarly, jumps have been classified into natural side jumps (N) or opposite side jumps (O).

The dataset was assembled by consulting the FIFA and UEFA official reports, as well as videos of every match. However, missing or unclear videos allowed us to obtain full information for only 220 out of 237 penalties of our dataset.[14] The 237 observations involve 198 kickers and 35 goalkeepers. The number of observations on individual players is too small to allow us to analyse the behaviour of single kickers (specifically, no single player in the sample kicks more than three penalties). On the other hand, the composition of this sample guarantees a relevant heterogeneity in terms of players' ability, because non-specialists as well as specialists are called to kick in shootout series.[15]

3.1 Kicker's specialization

The presence of a relevant number of non-specialists among the kickers in our sample is first suggested by the comparison of the scoring probability of our shootout penalties (71.3%, see tab. 2) with the success rates in the samples of penalties during regular time proposed by Chiappori et al. [2002], Palacios-Huerta [2003] and Dohmen [2008] (75.0%, 80.1% and 74.3%, respectively).[16] Our figures are more similar to the scoring probability in Apesteguia and Palacios-Huerta [2010] (73.1%), also obtained from a dataset composed of shootout penalties. An alternative interpretation of the lower performance of the kickers during shootouts could be associated to the higher psychological pressure of the players in this context.

In this section we try to understand whether the lower success rate of the kickers during shootouts has to be attributed to the presence of non-specialists or to performance decrements induced by psychological pressure. For this aim, we first introduce a specialization

[13] Consequently, we have assigned a left or right side even to kicks that in other papers would have been classified as 'central'. Thus, we avoid a discretionary width of the central area of the goal to classify kicks as central. In some cases, it is difficult to determine the precise position of the penalty, but this difficulty would have been the same (or worse) with other conventions.

[14] We were not able to find the video for eight out of twelve penalties of the match Czech Rep. vs. France at the UEFA Euro Cup 1996. We know that all of them were scored and the identity of the players, but we cannot classify the underlying actions. Additionally, the videos of nine other penalties were taken from an unfortunate perspective which made it impossible to identify the half of the goal area where they were directed. These penalties were obviously rather 'central'.

[15] In particular, we are thinking of the heterogeneity of kickers, because in a shootout series the goalkeeper remains the same (i.e., they are all specialists), while there must be at least five kickers for every team, so that even non-specialists must kick.

[16] The mean difference is significant ($p = 0.001$) only for the sample of Palacios-Huerta [2003].

variable. We define as 'specialist' a kicker who scored at least one penalty kick in his club (in league matches) or in his national team (excluding shootout penalties) during a four years interval centred around the observation date.[17] [18]
The explanatory power of our variable is confirmed by performance measures. As illustrated in tab. 3, we have identified 119 penalties kicked by specialists and 118 penalties kicked by non-specialists; the scoring probability in the first group is 76.5%, while the scoring probability in the second group is 66.1%. This difference is statistically significant, with $p = 0.039$. In the following discussion we will use the term 'specialist' and 'non-specialist' to identify players from these two sub-samples.
Apesteguia and Palacios-Huerta [2010] provide strong empirical evidence of the detrimental effects on performance of the 'importance' of the kick, as measured in terms of his weight on the ex-ante probability of winning the shootout. This evidence determines that:

- the kickers of the teams that are lagging in the score present lower scoring performance, and vice versa (a 'behind-ahead asymmetry');
- the detrimental effects on performance become more pronounced as the final rounds are approached.

Linked to these findings is the common wisdom that specialist players are usually al-

[17]Penalty statistics have been assembled from various public databases, in particular `www.transfermarkt.de` (for national team penalties) and `www.worldfootball.net` (for club penalties). We were able to obtain complete information concerning all penalty takers in the national team, while in 34 (out of 237) cases penalty scorers in the club were not available in the whole four years period. Notice that available statistics include only penalty scored (not kicked) during league matches, thus introducing a (negligible) endogenity issue. In all these cases we implicitly took a conservative approach, by assuming that no penalties have been kicked in absence of the information. In this sense, the category of non-specialists could include some specialist (and not vice versa). We also tried different alternative specifications for this variable, taking into consideration the general goal-scoring attitude of the players or collecting the penalties kicked during other tournaments (in particular, the continental club cups). Results are not significantly different.

[18]For a better illustration of the methodology used for building the specialization variable, an example can be useful. Germany vs. Argentina on June 30, 2006, the quarter-final of the FIFA World Cup, ended after a shootout. Eight penalties were kicked in the series, all scored for Germany, only two scored for Argentina. Germany consequently went through to the semifinals. The kickers for Germany were Ballack, Neuville, Podolski and Borowski. The kickers for Argentina were Cruz, Rodriguez, Cambiasso and Ayala. The observation period for assessing the specialization of the kickers is, as said, July 2004 - June 2008. During this period, only Ballack - out of the eight kickers - scored at least one penalty (three, in 2005, excluding shootouts) for his national team. During the same interval he also scored one penalty kick for his club team, Bayern München, in the 2004/05 season. Cruz during the same period never scored a penalty for Argentina, but he scored five penalties for his club team, Inter Milan (one, three and one, respectively, during the 2004/05, 2005/06 and 2007/08 seasons). Podolski, Neuville and Rodriguez never scored a penalty for their national team, but they scored respectively eight, eight and two penalty kicks for their club teams. All of these players were consequently classified as 'specialists', while Borowski, Cambiasso and Ayala never scored a penalty kick for their national or club team during those four years and were then considered as 'non-specialists'.

	Penalties kicked (a+b+c)	Goals (%) (a)	Penalties saved (% of failed kicks) (b)	Penalties out (% of failed kicks) (c)
Specialists	119	91 (76.5 %)	15 (53.6 %)	13 (46.4 %)
Non-specialists	118	78 (66.1 %)	28 (70.0 %)	12 (30.0 %)
Total	237	169 (71.3 %)	43 (63.2 %)	25 (36.8 %)

Table 3: The distribution of shot failures between specialists and non-specialists

located 'early' kicks in the shootout series, even if this choice is not always optimal.[19] Our classification of players' specialization supports such common wisdom: as reported in tab. 4, more than three quarters of the kickers in the first round are specialists, while their incidence falls to about one half and then one quarter in the following rounds.
One can then wonder whether specialists perform better because they are actually more talented, or because they simply kick 'less important' penalties.[20]
We have consequently tried to examine more in detail the effect of the 'importance' of every single penalty, and to verify whether specialists actually perform better than non-specialists also when psychological pressure is similar. Following the insight of Apesteguia and Palacios-Huerta [2010] concerning the 'behind-ahead asymmetry', we have then compared the scoring performance of specialists and non-specialists in the three cases when the player is called to kick when his team is ahead, even or behind in the score. Results are in tab. 5. As expected, the scoring performance is the highest when the team of the kicker is ahead in the score and the lowest when is lagging in the score. Interestingly, in every different situation, specialists always perform better than non-specialists.[21] We interpret this result as a confirm that specialists - as defined by our classification - are actually more talented than non-specialists.

4 A model of heterogeneous behaviour of the kickers

4.1 Two actions are not enough

The analysis of the data presented in the previous section is driven by the purpose of highlighting the role of the heterogeneous quality of the players. We are interested to understand how kicking ability affects the strategies chosen by professional players, and in

[19] Specialized players could be used early in order to obtain a greater chance to put pressure on the rivals, but it could also be rational to exploit their talent in the (possible) more 'important' final rounds.

[20] We thank a referee for attracting our attention to this fact.

[21] The mean difference is actually significant at the 1% level only when the team of the kicker is ahead in the score.

Round	Kickers	of which specialists (%)	Goals (%)
1	50	38 (76.0 %)	36 (72.0 %)
2	50	26 (52.0 %)	37 (74.0 %)
3	50	25 (50.0 %)	37 (74.0 %)
4	48	13 (27.1 %)	28 (58.3 %)
5	27	14 (51.9 %)	24 (88.9 %)
+5	12	3 (25.0 %)	7 (58.3 %)
Total	237	119 (50.2 %)	169 (71.3 %)

Table 4: Specialists distribution and scoring performance in the different rounds of the shootouts

particular whether two (or three) actions concerning the direction of the shot are sufficient to illustrate their different behaviour.
A first puzzling evidence is obtained by analyzing 'how' shots are failed. As shown in tab. 2, we see that 68 out of 237 penalties failed. Of these penalties, approximately two thirds of them (43) were cleared by the goalkeeper. The remaining (25) missed the goal area. In tab. 3 we illustrate how this proportion varies among specialists and non-specialists. Tab. 3 shows that specialists miss the goal area almost half of the times when they fail to score, while non-specialists miss the area only in 30% of the cases (difference is statistically significant with $p = 0.083$). Because specialists demonstrate superior skill (their overall performance is significantly higher), this empirical evidence is hard to understand within the framework proposed in the literature.
A different explanation for this puzzle might be the possibility that the side of the shot is not sufficient for adequately describing the strategic options of the kickers. In particular, each kicker must also select the difficulty level of his shot - for example, a kick that is more or less powerful, or closer to the post or to the centre.
Why would a kicker make his task more complicated? On the one hand, a kick close to the post (or very powerful) performs better - ceteris paribus - if the goalkeeper correctly guesses the side; on the other hand, a 'difficult' kick is more likely to miss the goal area. The ability to direct difficult kicks *within* the goal area is a trait of high-quality kickers. If this is the strategic framework, only two actions in the set of the kicker are not sufficient for illustrating the behaviour of the players.

Team of the kicker	Penalties kicked	Goals (%)	Specialists	Goals (%)	Non-specialists	Goals (%)
ahead	31	25 (80.6 %)	15	15 (100.0 %)	16	10 (62.5 %)
even	111	80 (72.1 %)	57	43 (75.4 %)	54	37 (68.5 %)
behind	95	64 (67.4 %)	47	33 (70.2 %)	48	31 (64.6 %)
Total	237	169 (71.3 %)	119	91 (76.5 %)	118	78 (66.1 %)

Table 5: Success rate for specialists and non-specialists in different score situations

When more than two actions are available to the players, the analysis becomes more complicated as it becomes possible that different types of players actually randomise over different actions. Furthermore, if there are more than two actions of the kicker to be analysed, representing the outcome of the game within a 2×2 matrix obtains wrong, or at least ambiguous, results.
In the next section we develop an extended version of the penalty kick game. In the new representation, the set of actions available to the kicker is composed of more strategic options than are usually analysed in the literature.

4.2 More options for kicking a penalty

In our version of the penalty game, the goalkeeper still decides only the side of his jump ((N)atural or (O)pposite), while the kicker decides the side *and* the difficulty level of the kick. It is not important here what makes a kick difficult; for our purpose, a (S)afe kick simply obtains a better payoff for the kicker than a (D)ifficult kick when players choose different sides and a worse payoff when they choose the same side.[22] The kicker, then, has four options (difficult/safe and natural/opposite). The normal form of the game is represented in tab. 6.
In line with our definition of a kick's difficulty, the action of the goalkeeper will be more crucial for S-type kicks. Formally, this means that $\pi_{iD} - P_{iD} \ll \pi_{iS} - P_{iS}$ $(i = N, O)$). π_{iD} is not much lower than the complement to 1 of the probability of missing the goal. Moreover, it is assumed that:

- the scoring probability is lower if the goalkeeper, ceteris paribus, correctly guesses the side (that is $\pi_{ij} > P_{ij}$ $(i = N, O;\ j = D, S)$;

[22] A safe kick, for example, has limited power and angle, but it is substantially always directed within the goal: if the goalkeeper correctly guesses the side, the probability of saving the goal is very high. On the contrary, a difficult kick can more easily miss the goal or hit the post/bar; if not, it is quite difficult to be blocked even if the goalkeeper jumps on the right side.

		Goalkeeper (G) N	Goalkeeper (G) O
Kicker (K)	ND	P_{ND}	π_{ND}
	NS	P_{NS}	π_{NS}
	OS	π_{OS}	P_{OS}
	OD	π_{OD}	P_{OD}

Table 6: Our extended version of the penalty kick game

- the scoring probability is higher, ceteris paribus, on the natural side of the kicker (that is $\pi_{Nj} > \pi_{Oj}$ and $P_{Nj} > P_{Oj}$).

Consequently, the ranking of the values assumed by the scoring is as follows:

$$\pi_{NS} > \pi_{OS} > \pi_{ND} > P_{ND}; \pi_{OD} > P_{OD} > P_{NS} > P_{OS} \tag{1}$$

The relative ranking between π_{OD} and P_{ND} is undetermined: it depends on the relative size of the advantage of a difficult kick when the goalkeeper correctly guesses the side with respect to the advantage obtained by kicking on the natural side. Notice that if $\pi_{OD} < P_{ND}$, the ND strategy dominates the OD strategy.

Being strictly competitive, this game obviously presents no pure strategy Nash equilibria. The characterisation of mixed strategy equilibria asks for the analysis of every case where both K and G randomise over two actions. Moreover, at the equilibrium, both players randomise over one kick/jump to the natural side and one to the opposite side. Hence, there are four strategy profiles to be analysed, as listed in tab. 7.

Kicker				Goalkeeper		
$prob(ND)$	$prob(NS)$	$prob(OS)$	$prob(OD)$	$prob(N)$	$prob(O)$	Profile
p_1	0	$1-p_1$	0	q_1	$1-q_1$	I
0	p_2	$1-p_2$	0	q_2	$1-q_2$	II
p_3	0	0	$1-p_3$	q_3	$1-q_3$	III
0	p_4	0	$1-p_4$	q_4	$1-q_4$	IV

Table 7: Possible equilibrium profiles

The profile that corresponds to the equilibrium of the game obviously depends on the actual values of the scoring probabilities, which in turn depend on the skills of the players. To make results more intuitive, we now introduce a simplifying restriction on the parameters. In particular, we fix $P_{iS} = 0$ and $\pi_{iS} = 1$. Such restrictions do not alter the nature of the results[23] and, moreover, are rather close to the empirical evidence. In other words,

[23]The characterisations of mixed strategy equilibria in the general case are available upon request.

this assumption states that if the difficulty level of the kick is low, i) the side advantage disappears, ii) the quality of the players does not matter, iii) the probability of missing the goal is null, and iv) should the goalkeeper guess correctly, he is almost sure to block the ball.[24]

The conditions for every single profile reported to be an equilibrium are detailed in Appendix A. The summary of the results is reported in tab. 8.

Profiles	I	II	III
p_{ND}	$\frac{1}{1+\pi_{ND}-P_{ND}}$	0	$\frac{\pi_{OD}-P_{OD}}{\pi_{ND}+\pi_{OD}-P_{ND}-P_{OD}}$
p_{NS}	0	$\frac{1}{2}$	0
p_{OS}	$\frac{\pi_{ND}-P_{ND}}{1+\pi_{ND}-P_{ND}}$	$\frac{1}{2}$	0
p_{OD}	0	0	$\frac{\pi_{ND}-P_{ND}}{\pi_{OD}+\pi_{ND}-P_{OD}-P_{ND}}$
q_N	$\frac{\pi_{ND}}{1+\pi_{ND}-P_{ND}}$	$\frac{1}{2}$	$\frac{\pi_{ND}-P_{OD}}{\pi_{ND}+\pi_{OD}-P_{ND}-P_{OD}}$
q_O	$\frac{1-P_{ND}}{1+\pi_{ND}-P_{ND}}$	$\frac{1}{2}$	$\frac{\pi_{OD}-P_{ND}}{\pi_{ND}+\pi_{OD}-P_{ND}-P_{OD}}$
Equilibrium conditions	$\pi_{ND}+P_{ND}>1$ and $\pi_{ND}(1-\pi_{OD}) \geq P_{OD}(1-P_{ND})$	$\pi_{ND}+P_{ND} \leq 1$	$\pi_{ND}+P_{ND}>1$ and $\pi_{ND}(1-\pi_{OD}) < P_{OD}(1-P_{ND})$

Table 8: Equilibrium profiles

These results reveal that only three types of equilibria are possible.
If the kicker is characterised by limited kicking skills (profile II; $\pi_{ND} + P_{ND} < 1$) the rational choice is to rely on safe kicks, that will never miss the goal area, hoping to wrong-foot the goalkeeper. Because in this case we do not model a natural side advantage, randomisation equals the probability of kicking left or right.
In the case of talented kickers (profiles I and III; $\pi_{ND} + P_{ND} > 1$) we have two sub cases: the equilibrium conditions distinguish when $\pi_{ND}(1 - \pi_{OD})$ is greater or lower than $P_{OD}(1 - P_{ND})$.
A lower value for $\pi_{ND}(1 - \pi_{OD})$ means that the kicker's ability to perform 'difficult' kicks regardless the side-choice of the goalkeeper is prevalent over natural side advantages

[24] Because we do not claim to be able to empirically distinguish between difficult and safe kicks, a direct confirm to our reasonable assumption is limited to π's. In this case, consider that in our entire dataset, when the goalkeeper was wrong-footed, he never saved the ball.

(profile III).[25] If this is the case, kickers will prefer to make difficult kicks on both sides. On the contrary, when $\pi_{ND}(1-\pi_{OD})$ is higher, the natural side advantage prevails (profile I). If this is the case, kickers will prefer to attempt difficult shots on the natural side and sometimes shoot safe on the opposite side to make their actions less predictable.

In conclusion, if we imagine a generic distribution of the three types of players in our sample, we expect that only talented players (that is, players with a sufficient ability to kick difficult) will try difficult kicks and that difficult kicks will be much more frequent on the natural side.

4.3 Empirical evidence about different behaviours of differently talented kickers

In this section we illustrate the empirical evidence drawn from our sample in support of the theoretical predictions proposed in the previous section. First, in tab. 9, we report the strategies chosen by the players in every penalty and the related outcome in terms of scoring frequencies. Because we are not able to distinguish between safe and difficult kicks, each cell aggregates the shots of both types. Thus, a significant phenomenon of aggregation bias may affect results. An evidence of aggregation bias is offered, for example, by considering that in tab. 9 $\pi_N < \pi_O$. This contradicts the widely accepted assumption of a natural side advantage.[26] If we assume that players behave as predicted by equilibrium mixed strategies, further evidence of aggregation bias is offered by violations of the properties of the Fundamental Lemma. In particular, the payoff of the goalkeeper associated with the pure strategy N ($1 - 73.8\%$) is lower than his payoff associated with the pure strategy O ($1 - 63.5\%$) with $p = 0.092$.[27] A more clear evidence will be proposed below, when we will separately analyze specialists and non-specialists
As for the elements in favour of the hypothesis of rational behaviour of the players, as discussed in sec. 2, Chiappori et al. [2002] obtain several properties of the equilibrium that need to be verified, even when aggregating the observations in a sample of heterogeneous players. The only condition for these properties to be true is that the ranking of the parameters must hold for every single player. The properties that survive in our setting predict that (i) the number of jumps to the natural side will be larger than the number of jumps to the opposite side, (ii) the number of kicks to the natural side will be larger than the number of jumps to the opposite side, and (iii) the profile (N,N) (i.e., the kicker chooses N and the goalkeeper chooses N) will be more frequent than the profile (O,O).[28] As observed in tab. 9, every property is confirmed in our sample. Testing the null hypothesis of equal propensities for each property leads to rejection at the 5-percent

[25] If no side advantage is in place, then $\pi_{ND} = \pi_{OD} > P_{ND} = P_{OD}$ and $\pi_{ND}(1-\pi_{OD}) < P_{OD}(1-P_{ND})$ because $\pi_{ND} + P_{ND} > 1$.

[26] See again footnote 8. This difference is actually not significant at the 10% level.

[27] Based on Pearsons χ^2 goodness-of-fit test of equality of two distributions (see Palacios-Huerta [2003, page 403] for details).

[28] Actually, those properties are verified for profiles I and III, while profile II predicts equal number of shots and jumps on natural and opposite side (see tab. 8).

		Goalkeeper (G)				
		Natural	Opposite	No move	No info	Total
Kicker (K)	Natural	72 (61.1%)	45 (84.4%)	2 (100.0%)	– (–)	119 (70.6%)
	Opposite	49 (91.8%)	47 (40.4%)	5 (100.0%)	– (–)	101 (68.3%)
	No info	5 (80.0%)	4 (100.0%)	– (–)	8 (100.0%)	17 (94.1%)
	Total	126 (73.8%)	96 (63.5%)	7 (100.0%)	8 (100.0%)	237 (71.3%)

Table 9: Players' behaviour and outcomes in the whole dataset: the number of observations and scoring probability (in parentheses)

level. Notice that the empirical evidence also confirms the natural side advantage when the goalkeeper correctly guesses the side of the kick (the scoring probability of (N,N) is significantly higher than the scoring probability of (O,O), $p = 0.014$).
These results are confirmed within each of the sub-samples of the specialists (see tab. 10) and non-specialists (see tab. 11). We expect lower heterogeneity (more limited effects of aggregation bias) in the first one because we cannot exclude the presence of some talented players in the sub-sample of non-specialists.
The higher talent of specialists is confirmed for every cell of the game.[29] The natural side advantage is particularly relevant for low quality players, which is in line with theoretical predictions. Interestingly, in tab. 10 (where heterogeneity among players is fairly limited) the expected payoffs between pure strategies (Fundamental Lemma) are very similar.[30]

Another interesting result concerns the predictability of the players' behaviour. Baumann et al. [2011], as discussed in sec. 2, offer empirical evidence to suggest that high-quality kickers (as defined by a ranking provided by a specialised magazine) are more predictable, in that they kick more often on the natural side; nonetheless, their scoring performance proves to be higher. Our empirical evidence confirms this result (specialist kickers kick approximately 56% of their shots on the natural side, while non-specialists kick approximately 52% of their shots on the natural side; the difference is actually not statistically significant).
Our conceptual model supports these findings. If we define a kicker's talent as the ability to score independently from the action of the goalkeeper, given a side advantage, player quality is associated with high values of P_{ND} and with low values of $\pi_{ND} - P_{ND}$. Given this definition of quality, the probability of kicking on the natural side (see again tab. 8,

[29]However, the difference is significant ($p = 0.014$) only for the action profile (O,O).

[30]See again footnote 27. The hypothesis that the scoring probability is identical across strategies for the kickers and the goalkeeper cannot be rejected at the 10% level of significance.

		Goalkeeper (G)				
		Natural	Opposite	No move	No info	Total
Kicker (K)	Natural	33 (63.6%)	27 (85.2%)	1 (100, 0%)	– (–)	61 (73.8%)
	Opposite	23 (95.7%)	23 (56.5%)	2 (100.0%)	– (–)	48 (77.1%)
	No info	3 (66.7%)	4 (100.0%)	– (–)	3 (100.0%)	10 (90.0%)
	Total	59 (76.3%)	54 (74.1%)	3 (100.0%)	3 (100.0%)	119 (76.5%)

Table 10: Specialists behaviour and outcomes: the number of observations and scoring probability (in parentheses)

		Goalkeeper (G)				
		Natural	Opposite	No move	No info	Total
Kicker (K)	Natural	39 (59.0%)	18 (83.3%)	1 (100.0%)	– (–)	58 (67.2%)
	Opposite	26 (88.5%)	24 (25.0%)	3 (100.0%)	– (–)	53 (60.4%)
	No info	2 (100.0%)	– (–)	– (–)	5 (100.0%)	7 (100.0%)
	Total	67 (71.6%)	42 (50.0%)	4 (100.0%)	5 (100.0%)	118 (66.1%)

Table 11: Non-specialists behaviour and outcomes: the number of observations and scoring probability (in parentheses)

		Goalkeeper (G) Natural	Opposite	Total
Kicker (K)	Natural	8 (11.1%)	7 (15.6%)	15 (12.6%)
	Opposite	4 (8.2%)	5 (10.6%)	9 (8.9%)
	Total	12 (9.5%)	12 (12.5%)	24 (10.1%)

Table 12: The distribution of out-of-the-goal shots among action profiles (frequency over the whole number of penalties in parentheses)

profiles I and III) actually increases when the kicker's quality increases.[31] This result is also confirmed by the empirical evidence provided by Baumann et al. [2011]. Notice that it strictly depends on the asymmetry of the game.
Of course, the term 'predictability', as in the paper by Baumann et al. [2011], has to be understood within a context of simultaneous moves. In this sense, we tried to verify whether this simultaneity assumption holds by analysing the contingency tables associated with strategies N and O for both specialists and non-specialists. As for the specialists, the statistical dependence can be rejected with $p = 0.61$, while in the subsample of non-specialists, statistical dependence can be rejected with $p = 0.08$, implying some ability of the goalkeeper to anticipate the move of the kicker. This evidence suggests another way to define ability in kicking penalties and produces results, as shown in tab. 11, that are less robust than the ones observed in tab. 10.

The results presented so far in this section confirm some properties of professional soccer player behaviour already obtained in similar previous applications and demonstrate that in more heterogeneous samples the effects of aggregation bias are increased. We want to conclude this section adding further direct support to the conceptual model presented in sec. 4.2. To this aim, we have analysed the distribution of shots outside-of-the-goal, i.e., shots that miss the goal or hit the post/bar. The predictions of our model in this direction are very clear: out-of-the-goal shots must be more frequent among specialists and on the natural side. Both predictions go against intuition, which might expect such major mistakes on the unfavoured side and from less talented players.
As for the distribution of out-of-the-goal shots between specialists and not specialists, see again tab. 3. The distribution of out-of-the-goal shots between natural and opposite side is presented in tab. 12. In both cases, the evidence confirms the predictions of our conceptual model.

[31] Imagine a very talented kicker who never fails the goal area even when shooting with great power and angle, especially on the natural side. This kicker is characterised by $\pi_{ND} = 1$ and $P_{ND} \to 1$ and will always kick difficult on the natural side; thus, his action will be easily predicted by the goalkeeper. This correct prediction will be of no use for the goalkeeper because the scoring frequency will be close to one.

5 Concluding remarks

In this paper, we have proposed an analysis of the behaviour of professional soccer players when kicking a penalty. With respect to the extant literature, to focus on the effects of heterogeneity that determines aggregation bias, we have proposed a more general game setting that expands the range of actions available to the kickers. The empirical implications of our model have been tested within an original dataset composed of every penalty kicked during shootouts at World Cup and Euro Cup since 1994.

First, in line with the literature, we find that the players in our sample - especially the high-quality ones - behave rationally, meaning that they move as suggested by game-theoretic predictions. Moreover, the observed outcomes support the framework proposed in our model.

Second, we claim that players' behaviours cannot be completely understood unless more than one strategic variable (a shot's difficulty level in addition to its direction) is considered. The analysis of the strategic choices concerning shot's difficulty is essential to deal with aggregation bias problems because it is along this variable that players differentiate their strategies conditional upon their own quality.

Finally, as far as predictability of the strategies of players, our model confirms that more talented players are more predictable, as suggested by the empirical evidence in the literature. The intuition for such a result is quite straightforward: a quality player will favour his preferred action (the natural side, in our case) because the action of the rival is less crucial for him.

References

Jose Apesteguia and Ignacio Palacios-Huerta. Psychological pressure in competitive environments: evidence from a randomized natural experiment. *American Economic Review*, 100(5):2548–2564, December 2010.

Michael Bar-Eli and Ofer Azar. Penalty kicks in soccer: an empirical analysis of shooting strategies and goalkeepers' preferences. *Soccer and Society*, 10(2):183–191, March 2009.

Michael Bar-Eli, Ofer H. Azar, Ilana Ritov, Yael Keidar-Levin, and Galit Schein. Action bias among elite soccer goal keepers: the case of penalty kicks. *Journal of Economic Psychology*, 28:606–621, 2007.

Florian Baumann, Tim Friehe, and Michael Wedow. General ability and specialization: evidence from penalty kicks in soccer. *Journal of Sports Economics*, 12(1):81–105, 2011.

Pierre André Chiappori, Steven D. Levitt, and Timothy Goreclose. Testing mixed-strategy equilibria when players are heterogeneous: the case of penalty kicks in soccer. *American Economic Review*, 92(4):1138–1151, 2002.

Germán Coloma. Penalty kicks in soccer - an alternative methodology for testing mixed-strategy equilibria. *Journal of Sports Economics*, 8(5):530–545, 2007.

Thomas J. Dohmen. Do professionals choke under pressure? *Journal of Economic Behavior and Organization*, 65:636–653, 2008.

Shih-Hsun Hsu, Chen-Ying Huang, and Cheng-Tao Tang. Minimax play at Wimbledon: comment. *American Economic Review*, 97(1):517–523, March 2007.

Kenneth Kovash and Steven D. Levitt. Professionals do not play minimax: evidence from Major League Baseball and the National Football League. *NBER Working Papers*, (15347), September 2007.

Wolfgang Leininger and Axel Ockenfels. The penalty-duel and institutional design: is there a Neeskens-effect? *Ruhr Economic Papers*, (4), March 2007.

Clark Miller. *He always puts it to the right. A concise history of the penalty kick*. Edmundsbury Press Ltd., London: Victor Gollancz, St., 1998.

Gian Carlo Moschini. Nash equilibrium in strictly competitive games: live play in soccer. *Economic Letters*, 85:365–371, 2004.

Ignacio Palacios-Huerta. Professionals play minimax. *Review of Economic Studies*, 70: 395–415, 2003.

Mark Walker and John Wooders. Minimax play at Wimbledon. *American Economic Review*, 91(5):1521–1538, 2001.

A Appendix

For every strategy profile reported in the table below, the equilibrium conditions are obtained.

Kicker				Goalkeeper		
$prob(ND)$	$prob(NS)$	$prob(OS)$	$prob(OD)$	$prob(N)$	$prob(O)$	Profile
p_1	0	$1-p_1$	0	q_1	$1-q_1$	I
0	p_2	$1-p_2$	0	q_2	$1-q_2$	II
p_3	0	0	$1-p_3$	q_3	$1-q_3$	III
0	p_4	0	$1-p_4$	q_4	$1-q_4$	IV

A.1 *Profile I*

This strategy combination is an equilibrium if the following conditions are met:

- The actions over which player K randomises (ND and OS) determine equal expected payoffs. Thus,

  $\hat{E}_K = \hat{E}(\Pi_K(ND)) = \hat{E}(\Pi_K(OS)) \quad \rightarrow \quad q_1 P_{ND} + (1-q_1)\pi_{ND} = q_1.$

  Consequently,

$$q_1 = \hat{E}_K = \frac{\pi_{ND}}{1+\pi_{ND}-P_{ND}}. \tag{2}$$

  Moreover, we need that

  $\hat{E}_K \geq \hat{E}(\Pi_K(NS)) = 1-q_1 \quad \rightarrow \quad \pi_{ND}+P_{ND} \geq 1$

  and

  $\hat{E}_K \geq \hat{E}(\Pi_K(OD)) = q_1\pi_{OD} + (1-q_1)P_{OD} \quad \rightarrow \quad \pi_{ND}(1-\pi_{OD}) \geq P_{OD}(1-P_{ND}).$

- The actions over which player G randomises (N and O) determine equal expected payoffs; thus,

  $\hat{E}_G = \hat{E}(\Pi_G(N)) = \hat{E}(\Pi_G(O)) \quad \rightarrow \quad p_1(1-P_{ND}) = p_1(1-\pi_{ND}) + 1 - p_1$

  Consequently,

$$p_1 = \frac{1}{1+\pi_{ND}-P_{ND}} \tag{3}$$

  and

$$\hat{E}_G = \frac{1-P_{ND}}{1+\pi_{ND}-P_{ND}} = 1-\hat{E}_K. \tag{4}$$

A.2 *Profile II*

This strategy combination is an equilibrium if the following conditions are met:

- The actions over which player K randomises (NS and OS) determine equal expected payoffs. Thus,

  $\hat{E}_K = \hat{E}(\Pi_K(NS)) = \hat{E}(\Pi_K(OS)) \quad \rightarrow \quad 1 - q_2 = q_2.$

  Consequently,

$$q_2 = \hat{E}_K = \frac{1}{2}. \tag{5}$$

  Moreover, we need that

  $\hat{E}_K \geq \hat{E}(\Pi_K(ND)) = (P_{ND} + \pi_{ND})/2 \quad \rightarrow \quad \pi_{ND} + P_{ND} \leq 1$

  and

  $\hat{E}_K \geq \hat{E}(\Pi_K(OD)) = (P_{OD} + \pi_{OD})/2 \quad \rightarrow \quad \pi_{OD} + P_{OD} \leq 1.$

  If the first of these two conditions is true, the second one need also to be true, given the rank of the values of the parameters.

- The actions over which player G randomises (N and O) determine equal expected payoffs; thus

  $\hat{E}_G = \hat{E}(\Pi_G(N)) = \hat{E}(\Pi_G(O)) \quad \rightarrow \quad p_2 = 1 - p_2.$

  Consequently,

$$p_2 = \hat{E}_G = \frac{1}{2} = 1 - \hat{E}_K. \tag{6}$$

A.3 *Profile III*

First of all, notice that a necessary condition for the Profile III to be an equilibrium is that $\pi_{OD} > P_{ND}$; otherwise, OD is dominated by ND. The profile III is an equilibrium if:

- The actions over which player K randomises (ND and OD) determine equal expected payoffs; thus,

  $\hat{E}_K = \hat{E}(\Pi_K(ND)) = \hat{E}(\Pi_K(OD)) \quad \rightarrow \quad q_3 P_{ND} + (1-q_3)\pi_{ND} = q_3 \pi_{OD} + (1-q_3) P_{OD}.$

  Consequently, we obtain

$$q_3 = \frac{\pi_{ND} - P_{OD}}{\pi_{ND} + \pi_{OD} - P_{ND} - P_{OD}} \tag{7}$$

  and

$$\hat{E}_K = \frac{\pi_{ND}\pi_{OD} - P_{ND}P_{OD}}{\pi_{ND} + \pi_{OD} - P_{ND} - P_{OD}}. \tag{8}$$

Notice that $\pi_{OD} > P_{ND}$ implies that $q_3 < 1$. Moreover we need that

$\hat{E}_K \geq \hat{E}(\Pi_K(OS)) = q_3 \;\; \rightarrow \;\; \pi_{ND}(1 - \pi_{OD}) \leq P_{OD}(1 - P_{ND})$

and

$\hat{E}_K \geq \hat{E}(\Pi_K(NS)) = 1 - q_3 \;\; \rightarrow \;\; \pi_{OD}(1 - \pi_{ND}) \leq P_{ND}(1 - P_{OD})$.

Given the rank of the values of the parameters, $\pi_{OD} < \pi_{ND}$ implies that $\pi_{ND}(1 - \pi_{OD}) > \pi_{OD}(1 - \pi_{ND})$ and $P_{OD} < P_{ND}$ implies that $P_{ND}(1 - P_{OD}) > P_{OD}(1 - P_{ND})$. Thus, if $\hat{E}_K \geq \hat{E}(\Pi_K(OS))$, also $\hat{E}_K \geq \hat{E}(\Pi_K(NS))$ must be true. Note that $\pi_{OD} > P_{ND}$ is a necessary condition for $\hat{E}_K \geq \hat{E}(\Pi_K(OS))$. In addition consider that the condition $\pi_{DN} + P_{DN} < 1$ implies that $\pi_{ND}(1 - \pi_{OD}) > P_{OD}(1 - P_{ND})$ so that $\pi_{DN} + P_{DN} \geq 1$ is a necessary condition for making the Profile III an equilibrium.

- The actions over which player G randomises (N and O) determine equal expected payoffs; thus,

  $\hat{E}_G = \hat{E}(\Pi_G(N)) = \hat{E}(\Pi_G(O)) \;\; \rightarrow \;\; p_3(1 - P_{ND}) + (1 - p_3)(1 - \pi_{OD}) = p_3(1 - \pi_{ND}) + (1 - p_3)(1 - P_{OD})$.

  Consequently, we obtain

$$p_3 = \frac{\pi_{OD} - P_{OD}}{\pi_{OD} + \pi_{ND} - P_{OD} - P_{ND}} \tag{9}$$

  and

$$\hat{E}_G = \frac{\pi_{OD} + \pi_{ND} - P_{OD} - P_{ND} - \pi_{ND}\pi_{OD} + P_{ND}P_{OD}}{\pi_{OD} + \pi_{ND} - P_{OD} - P_{ND}} = 1 - \hat{E}_K. \tag{10}$$

A.4 ***Profile IV***

This strategy combination is an equilibrium if the actions over which player K randomises (NS and OD) determine equal expected payoffs; thus,

$\hat{E}_K = \hat{E}(\Pi_K(NS)) = \hat{E}(\Pi_K(OD)) \;\; \rightarrow \;\; 1 - q_4 = q_4\pi_{OD} + (1 - q_4)P_{OD}$.
Moreover, we need that

$\hat{E}_K \geq \hat{E}(\Pi_K(ND)) = q_4\pi_{ND} + (1 - q_4)P_{ND} \;\; \rightarrow \;\; \pi_{OD}(1 - P_{ND}) \geq \pi_{ND}(1 - P_{OD})$

and

$\hat{E}_K \geq \hat{E}(\Pi_K(OS)) = q_4 \;\; \rightarrow \;\; 1 - P_{OD} \leq \pi_{OD}$.

The first condition, given the assumptions on the values of the parameters, is always false because $\pi_{OD} < \pi_{ND}$ and $1 - P_{ND} < 1 - P_{OD}$, so that the Profile IV can never be an equilibrium.