Psychological pressure in competitive environments: Evidence from a randomized natural experiment: Comment, Second Version[*]

Martin G. Kocher

University of East Anglia and University of Innsbruck

Marc V. Lenz

University of Cologne

Matthias Sutter[#]

University of Innsbruck, University of Gothenburg, and IZA Bonn

Abstract: Apesteguia and Palacios-Huerta (APH, forthcoming) report for a sample of 129 soccer penalty shootouts from various seasons in ten different tournaments that teams kicking first win significantly more often than teams kicking second by a margin of 21 percentage points. Collecting data for 470 shootouts, including all of APH's shootouts, we cannot replicate their result. Teams kicking first win 53.4% of shootouts, which is not significantly different from the *a priori* expected 50%. Our finding implies that (1) APH's results are not generally robust; (2) using selective subsamples without a coherent criterion for data inclusion might lead to non-representative results.

JEL-classification: C93

Keywords: Tournament, first-mover advantage, psychological pressure, field experiment, soccer, penalty shootouts

This version: 30 August 2010

[*] We are particularly indebted to Gary Charness, David Cooper, Uri Gneezy, Lorenz Goette, Steven Levitt, Bradley Ruffle, John Wooders, two referees and Joel Sobel for excellent comments and to TK Ahn, Esther Blanco, Joel Clovis, René Cyranek, Simon Czermak, Haoran He, Theresa Krismer, Klaus Löw, Dominik Matzat, Tamas Meleghy, Gina Neff, Julius Pahlke, Sebastian Strasser, and Stefan Vetter for invaluable help in collecting the data. We also want to express our gratitude to a number of football association officials, club managers, club historians and club secretaries that sent us reports of specific shootouts that took place in the 1970ies and 1980ies.

[#] Corresponding author. Address: Department of Public Finance, University of Innsbruck, Universitaetsstrasse 15, A-6020 Innsbruck, Austria. e-mail: matthias.sutter@uibk.ac.at

A recent paper by Jose Apesteguia and Ignacio Palacios-Huerta (forthcoming) reports a systematic and significant first-mover advantage in soccer penalty shootouts. As Apesteguia and Palacios-Huerta (hereafter APH) convincingly argue, soccer penalty shootouts provide an important example for studying the consequences of a puzzling psychological bias in sequential tournaments. More specifically, they say that their "results provide support for a source of psychological pressure that has a detrimental effect on performance, and that is different from others such as high stakes, social pressure or peer pressure previously documented in the literature" (p. 15f. of the accepted manuscript).

APH have collected data on penalty shootouts from ten different national and international tournaments (World Cup and Continental Cups for national teams, European-level club competitions and national cups for clubs). Based on their collection of 129 shootouts, APH show that the first-kicking teams won in 78 cases (60.5%), and lost only in 51 cases (39.5%), thereby giving the first-kicking team an advantage of 21 percentage points. This very large difference is significant ($p = 0.022$; two-sided Binomial test), despite the *a priori* probability of both teams winning with a 50% chance, since the two teams alternate in shooting (see section 1 for details of the rules). APH ascribe this strong effect of sequential moves in a tournament to psychological pressure on the second-kicking teams, because they typically face an intermediate score in the shootout that is to their disadvantage when kicking.

The striking findings of APH are at odds with the results of an earlier publication of ours. In Martin G. Kocher, Marc V. Lenz and Matthias Sutter (2008) we examined the determinants of scoring in penalty shootouts in the German Soccer Cup tournament (*DFB-Pokal*), but our focus in that paper was not on a possible first-mover advantage. Nevertheless, for a preliminary data set of 95 penalty shootouts in the seasons from 1986/1987 to 2006/2007 we found that the first-kicking teams won in only 48.4% of cases. This was not significantly different from the expected 50%, and hence different from the more recent results of APH.

Intrigued by this inconsistency and interested in providing a robustness test of APH's results, we have collected a data set that constitutes a strict superset of APH's initial set of data with almost four times its size. We have restricted the time period covered to all shootouts before summer 2003, thus following the convincing argument of APH that only until then shootout data can be used for an unbiased investigation of a potential first-mover advantage (see section 1 below for the reason). The tournaments that we consider include tournaments for national teams (World Cup, European Championship, Copa America, Asian Cup, African Cup of Nations, Gold Cup), European-level club tournaments (UEFA-Champions League, UEFA-Cup, European Cup Winners' Cup) and national cup tournaments

from the top-3 countries in the current UEFA-ranking of nations according to their clubs' performance in international competitions (i.e., from Germany, England and Spain). The only tournament included in our sample, but not in APH, is the European Cup Winners' Cup (which was perhaps excluded in APH because this competition was terminated by the UEFA in 1999).[1]

In total, we have been able to identify the order in which teams took kicks for 470 penalty shootouts from 1970 to June 2003. We find that in 53.4% of cases the team kicking first won the shootout. This fraction is not significantly different from 50% (p = 0.15; two-sided Binomial test). Hence, we are unable to replicate the results of APH, leading us to conclude that their claim of a significant first-mover advantage is not generally robust.

The reason for the difference in results might be due to the different methodologies used in the two papers to collect data. As becomes clear from their Table 1, APH select subsamples of shootouts in the various tournaments, covering different time periods. While APH take into account (almost) all shootouts in tournaments of national teams, the time periods that they choose for the various club tournaments do not reveal a systematic pattern or rule for inclusion. For the Champions League and the UEFA Cup, they consider the years 2000-2003 plus all finals. For the Spanish Cup they use data from 1999 to 2003 and all prior finals. For the German cups (German Soccer Cup Competition and German *Ligapokal*, also called Supercup) they take the years 2001-2003 as well as all prior finals. Finally, for the English cup competitions their paper lacks information on the years being taken into account.

In contrast, we have collected data for the entire history of eleven (out of twelve) tournaments included in our dataset. Disregarding Spanish Cup-data for the moment, we have been able to find the relevant information for 441 out of 463 shootouts (i.e., 95.2%) that ever took place in these eleven tournaments from 1970 to June 2003. The 22 missing data points are due to unavailability of data on kicking orders mainly in the 1970ies and early 1980ies. Only for Spain we were not able to find more shootouts with reliable information on the kicking order than APH did. Hence, we simply took the 29 matches for Spain reported by them. It is probably important to add that APH's Spanish data run clearly against the null-hypothesis of a non-existent first-mover advantage (see section 2). While we think that taking a subsample that represents only about 10% of the full Spanish data (with 280 shootouts in the

[1] Note that the Asian Cup of Nations is used in APH, but only for statistical analyses regarding the time period after June 2003. Since there were also shootouts in the Asian Cup before this period we include the relevant data.

Copa del Rey from 1970 to June 2003) could be problematic because the sampling error might be stronger than in the other eleven tournaments where we collected data on 95.2% of all shootouts that ever took place, we use these 29 Spanish shootouts as they guarantee that our sample is a strict superset of APH's dataset. Working with a superset of the data makes the comparison between the two papers straightforward.

The rest of this comment is organized as follows. In section 1 we describe the data and the rules that determine which team kicks first in a penalty shootout and how the outcome of a shootout is determined. Section 2 presents our results, and section 3 concludes.

1. Rules and Data

Penalty shootouts were introduced by the world governing body of soccer, the Fédération Internationale de Football Association (FIFA) in 1970 to determine the winner in knock-out tournament games in which there was a tie between two teams after 90 minutes of regular time and 30 minutes of overtime. Before 1970, the winner in tied games in a knock-out tournament was determined by the draw of a lot or a replay.

The basic rules for a penalty shootout are as follows (for details see the official "Rules of the Game" at www.fifa.com): First, each team selects five players (out of the players on the pitch in the 120th minute). Second, teams kick in alternating order. Third, the shootout is terminated as soon as the number of penalties converted by one team cannot be matched by the other team. If – after both teams have taken five kicks – both have scored the same number of goals, teams continue kicking in the same alternating order until one team has scored one goal more than the other from the same number of kicks (i.e., in a sequential one-on-one competition). Each penalty kick during the shootout must be taken by a different player, and all eligible players must have taken a kick before any player can take a second kick. In principle, this sequence can go on infinitely. However, the longest shootout in our data set includes 26 kicks, i.e., 13 for each team.[2]

From 1970 until June 2003 the team that won the referee's toss of a coin before the shootout *had* to take the first kick. Such a rule constitutes an explicit randomization to

[2] A soccer penalty shootout belongs to a class of games that Mark Walker, John Wooders, and Rabah Amir (forthcoming) call binary Markov games. For a theoretical account of such games, see their paper. Penalty shootouts are a good example for studying mixed strategies in games (see Pierre-Andre Chiappori, Steven Levitt and Timothy Groseclose, 2002). Consequently, soccer players – because of their penalty shooting experience – have been used to examine mixed strategy play of professionals in the laboratory, yielding mixed evidence, though (see Palacios-Huerta and Oscar Volij, 2008; Levitt, John List and David Reiley, 2010; Wooders, 2010).

determine the beginning team.[3] In July 2003, the FIFA changed the rules slightly by giving the winner of the referee's toss of a coin the *option* of choosing whether to kick first or second. This choice option may give rise to endogeneity problems, and therefore we only use data for 1970 to June 2003 (i.e., the end of the season 2002/2003), thus following the approach of APH.

Table 1 lists the different tournaments for which we collected data on the kicking order in penalty shootouts. Tournaments [1] to [6] are for national teams, including the arguably most important tournaments in the world of soccer, the World Cup and the European Championship. If psychological pressure were particularly strong in penalty shootout situations, it should be strongest in these tournaments, as they attract the most public attention and media coverage, and they are of the utmost importance for the standing and career prospect of every kicker on the pitch.[4] Tournaments [7] to [9] are for clubs on the European level, covering the most prestigious European Champions League (and its predecessor, the European Champion Clubs' Cup), the European Cup Winners' Cup (terminated in 1999) and the UEFA-Cup (and its predecessor, the European Fairs Cup). Tournaments [10] to [12] are national competitions where clubs compete for a national cup title. In Germany, the Soccer Cup Competition (*DFB-Pokal*) is far more important than the Supercup (*Ligapokal*), but we include the Supercup's four shootouts since they have also been included in APH. For England we consider the *FA-Cup* – the oldest soccer club tournament in the world – the *League Cup* and the *Charity Shield* (with only three shootouts in the history of the latter). APH's Spanish data originate from the *Copa del Rey*.

Below Table 1 we list the sources of our data. Among these sources is the LexisNexis website (www.lexisnexis.com) through which we got access to newspaper archives. These archives were extremely helpful in obtaining the necessary information for the Gold Cup and the English cup competitions[5]. Literally, we read through hundreds of newspaper clips to determine the kicking order for roughly 100 shootouts in the English cup competitions and the

[3] In the terminology of Glenn Harrison and List (2004) this rule guarantees a truly randomized natural experiment.

[4] However, very high stakes may have detrimental effects on performance, as has been shown under controlled conditions in the laboratory. See, for instance, Dan Ariely, Uri Gneezy, George Loewenstein, and Nina Mazar (2009).

[5] We had, first, identified the FA- and League-Cup matches that were decided by a penalty shootout from Tony Brown (2006, 2007). However, the two books do not provide any information on kicking orders in the shootouts.

five shootouts in the Gold Cup. Column [A] of Table 1 reports the number of shootouts for which we know which team took the first kick and whether this team won or not. In parentheses we show the number of shootouts for which we miss the information on the kicking order. For tournaments [1] to [11] we have been able to verify the necessary data for 441 out of 463 shootouts (95.2%). For Spain (tournament [12]) we resort to the 29 shootout-data from APH that constitute 10.4% of the 280 shootouts that took place in the *Copa del Rey* before summer 2003.

2. Results

Column [B] in Table 1 presents the relative frequency with which teams that took the first kick in a shootout won it. This relative frequency varies considerably across tournaments, reaching a low of 33% for the European Championship and a high of 69% for the African Cup of Nations and the Spanish data from APH. Note that in the presumably most important tournaments for national teams, the World Cup and the European Championship, the relative frequency is below 50%. However, given the relatively low number of observations for these competitions, it is, of course, more informative to look at the aggregate picture.

Since we are most confident in the representativeness of the data for the first eleven tournaments shown in Table 1, we report the overall relative frequency and the *p*-value of a two-sided Binomial test in row "*Sum [1] to [11]*". We find that 52.4% of shootouts were won by the team taking the first kick and that this percentage is not significantly different from 50% ($p = 0.34$; $N = 441$). In fact, we do not find a significant first-mover advantage for any of these eleven tournaments, as column [C] in Table 1 shows. It is particularly important to note that the winning percentage of the first-kicking team in the three competitions with the largest number of observations (Germany, England, UEFA-Cup) is either exactly or very close to 50% and certainly far away from a significant first-mover advantage.[6] Only for APH's Spanish data we note that a Binomial test yields a *p*-value of 0.06 for the 29 shootouts. If we add the Spanish data to tournaments [1] to [11], the first-kicking teams won in 53.4% of the

[6] We are even slightly biasing our results against the null hypothesis: There are two shootout results that allow – given the rules in section 1 – to conclude without any doubt that the first shooting team won (i.e., 5:3 and 4:1). Even when we did not find newspaper or other evidence on the order of kicking of shootouts with either of these two results in the 1970ies and 1980ies, we used all of them in our analysis. Hence, we have a slight selection bias working against our hypothesis that there is no first-mover advantage in shootouts, making our null result even stronger.

470 shootouts. However, the result still clearly fails to reach significant levels ($p = 0.15$; see row "*Sum [1] to [12]*").

Since the European Cup Winners' Cup (tournament [8]) was not taken into account in APH, it seems appropriate to apply the test to a subset of the data in which tournament [8] is excluded – while keeping APH's Spanish data. The result is reported in the final row of Table 1, showing that the overall frequency of first teams winning is 52.7%, which is again insignificantly different from 50% ($p = 0.27$; $N = 438$). Note that tournaments [1] – [7] plus [9] – [12] still constitute a strict superset of APH's data. Splitting this superset into APH's data ($N = 129$; 60.5% winning percentage of first-kicking teams) and the *Rest* ($N = 309$; 49.5% winning percentage of first-kicking teams) we find that APH's subset is significantly different from the *Rest* ($p < 0.05$; χ^2-test).

Table 1 and Figure 1 about here

Of course, the difference between APH's dataset and our larger one might be driven by APH using relatively more recent data. If first-kicking teams were more likely to win in more recent years than in earlier years, this might explain the difference. Therefore, we present in Figure 1 the relative frequency with which first-kicking teams won in each season from 1970/1971 (denoted 1970 in Figure 1) to 2002/2003 (denoted 2002).[7] Here we consider only competitions [1] to [11]. We notice ups and downs in Figure 1, but no systematic pattern in either direction over time. According to Binomial tests, the relative frequency of first-kicking teams to win a shootout is not significantly different from 50% in any single year.

[7] A year X in Figure 1 covers the shootouts that took place from August in year X to July in year X+1. This is important to consider, for instance, for World Cups and European Championships that take place every four years, typically from June to early July. These competitions are then regarded as the end of a season. Therefore, the World Cup 2002, for example, is included in season 2001/2002, thus showing up in the bar "2001" in Figure 1. The number of shootouts is generally smaller in the 1970ies and 1980ies due to a change in rules of national cup competitions. In the late 1980ies and early 1990ies it was decided to proceed immediately to a penalty shootout if a match ended with a draw after 120 minutes, rather than continuing with the former tradition of staging a replay (or even more than one replay, if necessary, as in England). This new practice – introduced to save on the number of matches to be played – increased the number of shootouts considerably.

3. Conclusion

Penalty shootouts from 1970 until June 2003 constitute a truly randomized experiment since the order of taking penalty kicks was exogenously determined by a random device. Taking a look at the influence of the order of kicking on the probability of winning allows studying a possible first-mover advantage that might be due to psychological pressure on the team that kicks second. While Apesteguia and Palacios-Huerta (forthcoming) have reported a significant first-mover advantage in 129 shootouts sampled from ten different tournaments and different time periods, we have not been able to replicate their central finding. In fact, collecting a dataset of 470 shootouts that constitutes a strict superset of APH's data, we fail to find a significant first-mover advantage, despite the fact that our sample is almost fourfold the size of the one in APH.

We think that there are two main implications from this comment: First, the findings and conclusions offered by Apesteguia and Palacios-Huerta (forthcoming) on the first-mover advantage in sequential tournaments do not appear to be robust and, therefore, are less general than it seems at first sight. Second, taking specific subsamples without a clear and coherent criterion on which time periods and competitions to include can lead to non-representative results.

References

Apesteguia, Jose, and Ignacio Palacios-Huerta. Forthcoming. "Psychological Pressure in Competitive Environments: Evidence From a Randomized Natural Experiment." *American Economic Review*.

Ariely, Dan, Uri Gneezy, George Loewenstein, and Nina Mazar. 2009. "Large Stakes and Big Mistakes." *Review of Economic Studies* 76(2): 451-470.

Brown, Tony. 2006. *The FA Challenge Cup Complete Results. 1871/72 to 2005/06.* Nottingham: SoccerData Publication.

Brown, Tony. 2007. *The Football League Cup Complete Results. 1960/61 to 2006/07.* Nottingham: SoccerData Publication.

Chiappori, Pierre-Andre, Steven Levitt, and Timothy Groseclose. 2002. Testing Mixed-Strategy Equilibria When Players are Heterogeneous: The Case of Penalty Kicks in Soccer. *American Economic Review* 92(4): 1138-1151.

Harrison, Glenn, and John A. List. 2004. "Field Experiments." *Journal of Economic Literature* XLII(4): 1013-1059.

Ionescu, Romeo. 2003. *The Complete Results & Line-Ups of the European Fairs Cup 1955-1971.* Cleethorpes: Soccer Books Limited.

Ionescu, Romeo. 2004a. *The Complete Results & Line-Ups of the European Champion Clubs' Cup 1955-1991 – The Knockout Years.* Cleethorpes: Soccer Books Limited.

Ionescu, Romeo. 2004b. *The Complete Results & Line-Ups of the European Champions League 1991-2004.* Cleethorpes: Soccer Books Limited.

Ionescu, Romeo. 2004c. *The Complete Results & Line-Ups of the European Cup Winners' Cup 1960-1999.* Cleethorpes: Soccer Books Limited.

Ionescu, Romeo. 2004d. *The Complete Results & Line-Ups of the UEFA Cup 1991-2004.* Cleethorpes: Soccer Books Limited.

Ionescu, Romeo. 2004e. *The Complete Results & Line-Ups of the UEFA Cup 1971-1991.* Cleethorpes: Soccer Books Limited.

Ionescu, Romeo. 2010. *The Africa Cup of Nations 1957-2010. A Statistical Record.* Cleethorpes: Soccer Books Limited.

Kocher, Martin G., Marc V. Lenz, and Matthias Sutter. 2008. "Performance Under Pressure – The Case of Penalty Shootouts in Football." In: Patric Andersson, Peter Ayton, and Carsten Schmidt (eds.). *Myths and Facts about Football: The Economics and Psychology of the World's Greatest Sport*. Cambridge: Cambridge Scholars Publishing: 61-72.

Levitt, Steven, John List, and David Reiley. 2010. "What Happens in the Field Stays in the Field: Exploring Whether Professionals Play Minimax in Laboratory Experiments." *Econometrica* 78(4): 1413-1434.

Palacios-Huerta, Ignacio, and Oscar Volij. 2008. "Experientia Docet: Professionals Play Minimax in Laboratory Experiments." *Econometrica* 76(1): 71-115.

Walker, Marc, John Wooders, and Rabah Amir. forthcoming. "Equilibrium Play in Matches: Binary Markov Games." *Games and Economic Behavior*.

Wooders, John. 2010. "Does Experience Teach? Professionals and Minimax Play in the Lab." *Econometrica* 78(3): 1143-1154.

Table and Figure

Table 1: Shootout data from 1970/1971 to 2002/2003

Tournament	[A] Number of shootouts with order of kicks known (unknown)	[B] First-kicking team wins (relative frequency)	[C] p-value (two-sided Binomial test)
[1] World Cup	16 (0)	0.438	0.80
[2] European Championship	9 (0)	0.333	0.51
[3] Copa America (South America)	12 (0)	0.667	0.39
[4] African Nations Cup	13 (4)	0.692	0.27
[5] Gold Cup (Northern and Middle America)	5 (0)	0.400	1.00
[6] Asian Cup	8 (2)	0.375	0.73
[7] European Champions League*	28 (0)	0.643	0.19
[8] European Cup Winners' Cup[#]	32 (2)	0.625	0.22
[9] UEFA-Cup[‡]	74 (2)	0.527	0.73
[10] German Cups (DFB-Pokal, Ligapokal)	122 (1)	0.500	1.00
[11] English Cups (FA-Cup, League Cup, Charity Shield)	122 (11)	0.500	1.00
Sum [1] to [11]	*441 (22)*	*0.524*	*0.34*
[12] Spanish Cup (Copa del Rey; APH-data)	29 (251)	0.690	0.06
Sum [1] to [12]	*470 (273)*	*0.534*	*0.15*
Sum [1] to [12], except [8]	*438 (271)*	*0.527*	*0.27*

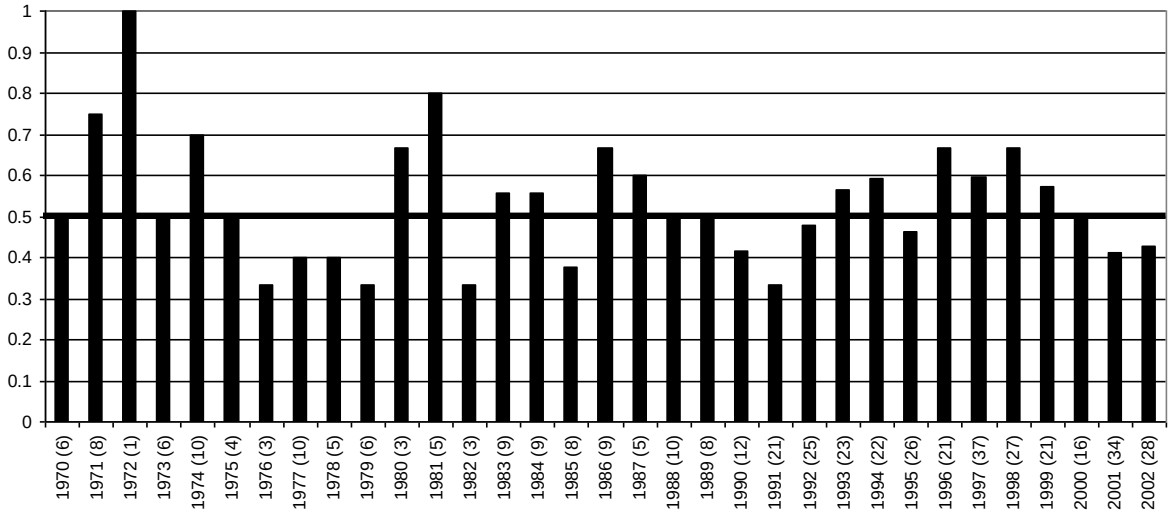
Sources: [1] www.fifa.com, www.kicker.de. [2] www.kicker.de, www.fussballdaten.de. [3] www.rsssf.com. [4] Romeo Ionescu (2010), *L'Equipe* (print edition), e-mails to football associations. [5] www.lexisnexis.com. [6] en.wikipedia.org, swe.worldfootball.net, sports.sina.com.cn, sports.sohu.com. [7] Ionescu (2004a, 2004b). [8] Ionescu (2004c). [9] Ionescu (2003, 2004d, 2004e). [10] *Kicker Sportmagazin* (print edition), www.kicker.de, www.fussballdaten.de, www.weltfussball.de, e-mails to clubs. [11] Brown (2006, 2007), www.lexisnexis.com, e-mails to clubs. [12] Apesteguia and Palacios-Huerta (forthcoming).

* Data include preceding European Champion Clubs' Cup, which was played out from 1955-1991.

[#] This tournament ceased in 1999. It is not included in the set of tournaments used in APH.

[‡] Data include preceding European Fairs Cup, which was played out from 1955-1971.

Figure 1: Relative frequency with which the first-kicking team wins a shootout in seasons 1970/71 to 2002/03



Note: x-axis indicates season (from August in year t to July in year t+1). Data are pooled for tournaments [1] to [11]. Numbers in parentheses indicate the number of shootouts in a particular season.