

Do Behavioral Considerations Cloud Penalty-Kick Location Optimization in Professional Soccer?

Classical/Statistical Game Theory & Empirical Testing using Polynomial and *ML Regularized Lasso* Regression

Paper Track: Soccer

Paper ID: 20251395

github.com/sdsander-syr/soccer_pk_shot_location

1. Introduction

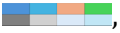
In soccer, penalty kicks (PKs) are taken with fair regularity (~ once in four matches [1]) and often constitute high-leverage, or game-pivotal, events given the sport's low-scoring nature. In a tabulation of 294,970 international, professional league, and professional cup match results recorded on footystats.org [2], we find that 1-0 and 1-1 are the most common professional full-time match scorelines, occurring 17.9 and 11 percent of the time, respectively. These outcomes are followed by 2-1 (8.5%) and 0-0 (7.7%). Across all recorded match outcomes, average full-time goals per professional match are 2.85. In the 2022-23 EPL, matches averaged 0.26 PKs and 0.194 PK-conversions, equivalent to about 6.8% of goals scored according to tabulations of data from transfermarket.com [3]. In 2023-24, this percentage rose to 7.6% [4]. Herein, we examine whether professional penalty-takers strictly optimize on expected conversion-rate when choosing shot-location, or whether behavioral considerations, such as "looking credible" by not missing the goal-space entirely, are also at play.

There is a famous sports adage that goes, "A win is a win." Despite the popularity of the adage, behavioral considerations often cloud player- and coaching-optimization in sport. Perceptions cause NFL coaches to invoke sub-optimal play-calling [5], which is not fully explained by risk-aversion [6]. Among others, notable NBA Centers Shaquille O'Neal (retired) and Andre Drummond (active) vowed never to shoot free-throws underhanded, given the shot's optics, despite demonstrated associated improvements for large-handed players [7],[8]. Both players in fact rate as abysmal free-throw shooters, with O'Neal shooting 52.7% for his career, and Drummond shooting 48.2%. By comparison, NBA players in 2023-24 averaged about 78.4% from the free throw line, with a standard deviation from player-to-player of approximately 9 percentage points. Therefore, the threshold for a negative outlier is 51.4%; Drummond's free throw accuracy is negatively-outlying by the 3 standard-deviation rule, whereas O'Neal's is close.

These low free-throw percentages spawned an opposing-team strategy of purposely and repeatedly fouling such players (e.g., the "Hack-a-Shaq" defense). Like an intentional walk in baseball, such a strategy automatically sent the player to the line to shoot free throws. Drummond once shot 36 free throws in a game, making 13, and O'Neal once shot 39 in a playoff game, making 18. In a league that has long-averaged ~1.1 points per possession, repeatedly giving 2 free throws to either player equated to strong defense, in expectation. In other words, their free-throw shooting was a problem for their respective teams. Be that as it may, O'Neal called the underhand free throw a "shot for sissies." For his part, Drummond tweeted, "Let me make this clear.... I'm not shooting free throws underhand." One might gather from this anecdote that certain players don't care about certain

important performance aspects. However, Drummond surely cares a great deal about his free-throw shooting, even going so far as to change his primary shooting hand, first in practice and then in games, in the middle of his NBA career. But he simply would not shoot underhanded. In a similar vein, NBA players shooting corner three-point attempts hit the front of the rim twice as often as they hit the side of the backboard. Given that the latter outcome is associated with novice shooting, players tend to err disproportionately in the other direction, decreasing their overall accuracy on the shot [9]. These examples suggest that professional athletes prefer to “keep up appearances” in playing style, even at the ironic cost of performance decrements. Herein, we test whether this behavior applies to penalty kicks in the *beautiful game*.

2. Methods

We study the spatial allocation of penalty kicks first game-theoretically, within a pair of modelled games — one classical and the other a statistical game of chance — and then by analyzing a large sample of professional PK-attempts empirically. In the game-theoretic analysis, we solve each game to develop empirical predictions of PK-shot location distribution. For the complementary empirical analysis, we use linear, polynomial, and ML-regularized Lasso regression models to further estimate properties of the observed PK shot-location distribution. As in Almeida et al. (2016) [10], we divide the goal-face planar-region into 8 partitions in the empirical analysis, to track shot-locations as the ball crosses or nearly crosses. The partition space is represented as , where each partition represents one-eighth of the total goal area (4 ft x 6 ft = 24 ft² per partition; 8 ft x 24 ft = 192 ft² for the total). Facing the goal and moving clockwise, from the top-left partition, we call these: {Top Left, Top Middle Left, Top Middle Right, Top Right, Bottom Right, Bottom Middle Right, Bottom Middle Left, Bottom Left}. Off-target penalty shots in the data are pinned to the Euclidean-nearest partition as being the shooter’s intended target. Moreover, professional shooters are taken as sufficiently skilled to always shoot to or nearest to their intended partition. With this spatial set-up, we study the relationship between shot-volume-to-partition and partition-conditional conversion-rate across penalty-shot partitions, for all (n = 536) penalty kicks in the men’s 2015-2020 UEFA Champions and Europa Leagues, the most elite levels of international club football in Europe.

3. Results

3.1 Classical Game-Theoretic Model: PK-Attempt as a Classical Game between PK-Taker and Goalkeeper

We first consider a discrete, constant-sum, one-shot penalty-kick game between two players: penalty-taker (T) and goalkeeper (G), or $N = \{T, G\}$. For expositional purposes, and unlike the empirical section to follow, we restrict or simplify the game-theoretic strategy sets for each player to an essential difference in shot locations: center partition or corner partition. The empirical analysis will simultaneously consider the decision between top and bottom partition. It can be easily verified that the succeeding results are not unique to this strategy set restriction. We represent T ’s strategy set as $S_T = \{\text{Shoot Center}, \text{Shoot Corner}\}$. That is, *Player T* can either shoot to the center or shoot to the corner. For *Player G*, the strategy set is $S_G = \{\text{Stay Center}, \text{Dive Corner}\}$. The payoff to *Player T*, π_T , is the conversion rate, whereas the payoff to *Player G*, π_G , is the non-conversion rate or the complement of the conversion rate. For the game, we select conversion rates that are ordinally consistent with the underlying PK-data, such that any equilibria found are qualitatively consistent with subsequent empirical equilibria. For a given

strategy profile or game trial (e.g., $S = \{S_T, S_G\} = \{\text{Shooter Center}, \text{Stay Center}\}$), then, we have that $\pi_T = 1 - \pi_G$. Lastly, the set of strategy profiles is simply the Cartesian product $S_G \times S_T$ such that there are $(2 \cdot 2)$ or 4 possible strategy profiles in the game. The following normal-form game representation of Table 1 summarizes payoffs and conditional best-response strategies in the game.

Table 1: Normal Form Representation of PK-game between PK-Taker and Goalkeeper

		Player G	
		Stay Center	Dive Corner
Player T	Shoot Center	$(\frac{3}{10}, \frac{7}{10})$	$(\frac{9}{10}, \frac{1}{10})$
	Shoot Corner	$(\frac{9}{10}, \frac{1}{10})$	$(\frac{7}{10}, \frac{3}{10})$

*Arrows represent best response given opponent's strategy.

Table 1 verifies, via expositional best-response arrows, that a pure strategy Nash equilibrium does not exist in the game. In *A Course in Game Theory* (MIT, 1994 [11]), Osborne and Rubinstein provide an essential definition of a Nash equilibrium as a point or strategy profile in a non-cooperative game at which no player can gain from unilateral deviation of strategy. Alternatively, a Nash equilibrium is a point or strategy profile of mutual best response. In the above normal-form game, each cell is a strategy profile. A pure strategy Nash equilibrium, then, is a cell for which all best-response arrows point toward that cell. It is straightforward to see that no such strategy profile exists in the above game. Of course, Nash (1950 [12]) proved in his seminal paper on non-cooperative game solutions that at least one (Nash) equilibrium must exist in pure or mixed strategy, a result that he developed from the generalized fixed-point theorem of Kakutani (1941 [13]). The implication of this result herein is that the present game possesses a *mixed strategy Nash equilibrium (msNe)*. The present game is a form of mixed coordination game, whereby T would like to avoid coordinating on shot location, and G would like to coordinate on shot location. It is typical for coordination games to have *msNe* solutions, as such games lack a focal point. Robustly, a *msNe* emerges given both expected payoff-maximizing players and given “minimax players,” or players who minimize their opponent’s maximum score. Under fixed-sum games such as this, it is a general result that the same *msNe* emerges whether players are payoff-maximizing or “maximin” in their objectives.

We solve for the *msNe* as the likelihood (p_t^*) that T chooses *Shoot Center* and the likelihood (q_G^*) G chooses *Stay Center*. Of course, these likelihoods imply their respective complements: the likelihood ($1 - p_t^*$) that T will choose *Shoot Corner* and the likelihood ($1 - q_G^*$) that G will choose *Dive Corner*. Then, a *msNe* is a set of probability allocations, $\{p_t^*, q_G^*\}$, such that the opponent is left indifferent between her available strategies. That is, p_t^* is the frequency of T choosing *Shoot Center* such that G is indifferent on the margin between choosing *Stay Center* and *Dive Corner*. Conversely, q_G^* is the frequency of G choosing *Stay Center* such that T is indifferent on the margin between choosing *Shoot Center* and *Shoot Corner*. Formally,

$$p_t^* = \{p_t \in [0,1]: \pi_G(\text{Stay Center}|p_t) = \pi_G(\text{Dive Corner}|p_t)\} \quad (1)$$

and

$$q_G^* = \{q_G \in [0,1]: \pi_T(\text{Shoot Center}|q_G) = \pi_T(\text{Shoot Corner}|q_G)\} \quad (2)$$

As such, we use Table 1 to obtain p_t^* as the solution to the equation:

$$7p_t + .1(1 - p_t) = .1p_t + .3(1 - p_t) \quad (3)$$

Then, $p_t^* = \frac{1}{4}$ and $(1 - p_t^*) = \frac{3}{4}$. Similarly, we obtain q_G^* as the solution to the equation:

$$3q_G + .9(1 - q_G) = .9q_G + .7(1 - q_G) \quad (4)$$

Then, $q_G^* = \frac{1}{4}$ and $(1 - q_G^*) = \frac{3}{4}$. Equilibrium payoffs are $\pi_T^* = \frac{3}{4}$ and $\pi_G^* = \frac{1}{4}$, which reflect the expectation of the observed shot conversion and miss rates (i.e., across all shot locations) under equilibrium. This corresponds closely to empirically-observed conversion-rates. It is a general result that a *msNe*, when it exists, creates expected **invariance** between strategies. It does so because players forming a mixed strategy are more likely to choose a strategy that features a higher expected payoff for them. However, this reliance is anticipated and countered by the opponent until an equilibrium emerges. In this case, G *Dives Corner* most often, in anticipation that T will *Shoot Corner* most often. This drives down returns from T choosing *Shoot Corner* to the point that expected return from *Shoot Corner* and *Shoot Center* are equalized, at $\frac{3}{4}$ in this case. This result is general to games with *msNe* and not specific to the payoff values selected herein. In fact, the result is definitionally-guaranteed under existence of a *msNe*.

This classical game-theoretic analysis has shown that a *msNE* exists in the PK-game. As such, classical game-theory predicts for the PK-game that there is no expected correlation, positive or negative, between location-conditional shot-volume and conversion-rate in the game's equilibrium. For the empirical section to follow, then, this game-theoretic finding predicts that, under optimal or equilibrium shot-location allocations, there will be no significant correlation in either direction between partition-level PK conversion-rate and partition-level PK shot-volume across any subsets

of the partitions. According to this theoretical result, any relationship between these variables will be driven away by strategic interaction if strategies are selected optimally.

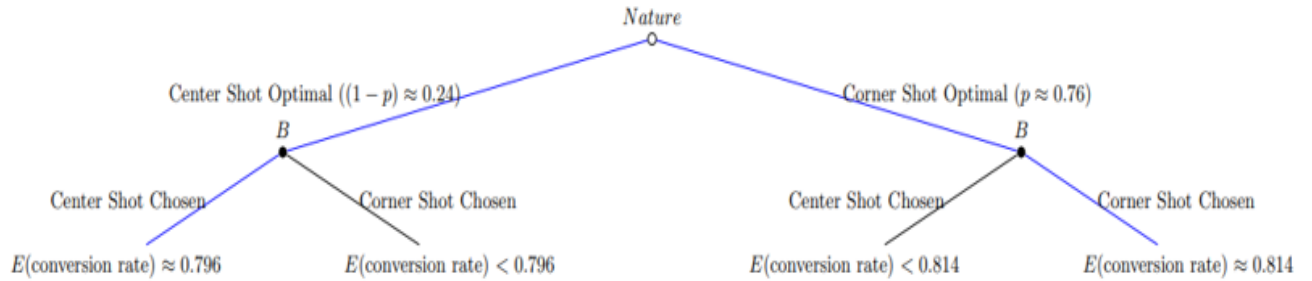
3.2 Alternative Theoretical Model: PK-Attempt as a Mixed (Continuous-Discrete) Statistical Game of Chance between PK-Taker and Nature

From a cursory examination of the study data to be analyzed, there emerges evidence that goalkeepers do not directly influence shot conversion-rates to all partitions, or do so fairly homogeneously. For example, observed shots to the top-right partition are *never* saved in the data. They either are converted or miss the goal entirely. In this sense, the present PK-game has properties of a statistical game of chance. Specifically, we can alternatively treat PK-attempts as a game between the PK-taker, T , and “nature” or “chance,” C . That is, $N = \{C, T\}$. In this alternative game, C first selects, from a continuous distribution, game conditions that influence the payoffs for different shot-location strategies by T . Observing each draw, T then selects a shot-location partition from the same discrete strategy set as in the classical game (i.e., $S_T = \{\text{Shoot Center}, \text{Shoot Corner}\}$). For example, the *a priori* expected conversion-rate of a Center shot and a Corner shot, respectively, are affected by *nature* from game-to-game, such as by *nature* presenting T with different weather conditions and different goalkeepers. In the statistical game, *nature* also selects how effective that goalkeeper is on a given night. Let us assume that, for the representative T in match i , $\pi_{T,i}(\text{Shoot Corner}) \sim N(\mu = 0.8, \sigma = 0.05)$ and $\pi_{T,i}(\text{Shoot Center}) \sim N(\mu = 0.75, \sigma = 0.05)$, and that these two expected payoffs for match i are independently drawn by C , according to the specified normal distributions, via the game/setting conditions selected by C . This independence assumption takes the conditions that influence shooting to the center as fundamentally different from those affecting shooting to the corner. For example, shooting to the corner relies heavily on (conditions to support) pinpoint accuracy, while shots to the center may rely more heavily on (conditions to support) a fast approach, a firm plant, and subsequent pace.

We further assume T has rational expectations as to how chance factors influence PK conversion-rates at the shot level. That is, T has informed, unbiased estimates of his or her shot-location and match conditional conversion-rates, $\pi_{T,i}(\text{Shoot Corner})$ and $\pi_{T,i}(\text{Shoot Center})$, before taking a given PK. Rational expectations is a microeconomic theory of decision-making that assumes agents use all available information to make unbiased predictions and optimal choices based on those predictions. It was developed by economists John Muth [14], Robert Lucas and Edward Prescott [15], and Thomas Sargent [16], where the latter three scholars won Nobel Prizes partly for contributions to the idea. Then, chance assigns location-conditional expected conversion-rates; T infers his or her match-conditional expected-payoff draws from observing conditions and the opposing goalkeeper, as well as from experience. Then, T receives the pair of expected payoffs after they are drawn independently by C , and T then selects and executes the highest expected payoff shot location for that shot attempt instance. We simulate these chance-based expected payoff pairs for 1 billion PK takes, finding that T selects *Shoot Corner* approximately 0.76 proportion of the time in this game. That is, $\pi_{T,i}(\text{Shoot Corner}) > \pi_{T,i}(\text{Shoot Center})$, with likelihood 0.76 across simulated match draws. This implies that T selects *Shoot Center* with likelihood $(1 - 0.76) = 0.24$. Though only calibrated with respect to the data and not with respect to the previous classical game-theoretic setting, it is interesting to note that these shot-volume likelihoods match closely to *msNe* likelihoods of shooting corner and shooting center for T . Moreover, $\pi_{T,i}(\text{Shoot Corner}) = 0.80000002$, as close to expected in the simulation, whereas $\pi_{T,i}(\text{Shoot Center}) = 0.75000023$, as also close to expected.

However, we are not interested in the overall conversion-rates for these shot-locations but the simulated location-dependent conversion-rate when the *PK* is actually shot to that location (i.e., when that location has the higher expected value such that it is the revealed or optimal shot location). The former rates are never observed, whereas the latter are. This is because *PK*-takers in the game only shoot to the optimal location for that shot (i.e., under the conditions drawn by chance). That is, we only observe conversion-rates at locations when those locations are optimal and, therefore, chosen for that trial. We find the latter observed/optimal conversion-rates within the statistical game to be $\pi_{T,i}^*(\text{Shoot Corner}) = 0.814$, as close to expected, in the simulation, whereas $\pi_{T,i}^*(\text{Shoot Center}) = 0.796$. When observed, Corner shots are converted at an equilibrium rate of 81.4% when taken in the statistical game; Center shots are converted at an observed equilibrium rate of 79.6% when taken. Figure 1 represents the game in extensive, or game-tree, form.

Figure 1: PK Statistical Game in Extensive Form



A game tree is a directed graph, where the parent node initializes the game, subsequent child nodes represent each possible game state, and terminal nodes represent outcomes. Edges from a given node represent possible actions from that node. In this case, *nature* chooses either the Center shot to be optimal for the shooter $((1 - p^*) \approx 0.24)$ or the Corner shot to be optimal $(p^* \approx 0.76)$. Then, the payoff-maximizing player selects either to *Shoot Corner* or *Shoot Center* given *nature's* draws for that trial. Blue paths from initial to terminal node represent the two possible equilibria for a given trial conditional on *nature's* draws for that trial.

From the statistical game results, there is a clear, positive correlation between location-conditional conversion-rate and shot-location in the aggregate data of the statistical game of chance simulation, whereby the 0.814 equilibrium conversion-rate location (*Corner*) is targeted with likelihood 0.76, and the 0.796 equilibrium conversion-rate location (*Center*) is targeted with likelihood 0.24 (across $n = 1$ billion simulated *PK*-shots).

Over this big-data simulation sample, we find a positive relationship between equilibrium shot-location frequency and equilibrium location-dependent conversion-rate. Simulated *PK*-shots were roughly thrice as likely to go to the relatively high equilibrium conversion-rate location in the game. Given that the simulation sample size is asymptotically large, we could take the relationship as significant on faith. To test for the relationship, however, we also create a sampling distribution of *PK*-simulations in *R*. We do so by resampling on the simulation code to form 100 samples, each with 100 independent statistical *PK*-game trials. Each trial has the same underlying properties as in the previous-simulated statistical game of chance. For each sample, we obtain mean (equilibrium)

conversion-rate for Corner shots when chosen, mean (equilibrium) conversion-rate for Center shots when chosen, proportional volume of Corner shots, and proportional volume of Center shots. This gives us 2×100 or 200 sample-statistic data-pairs of (shot-volume to location, conversion-rate at location when chosen). Figure 2 presents the game-generating PK-simulation and resampling code in R.

Figure 2: Statistical PK-Game of Chance Simulation and Resampling Code

```

{r}
#a is corner, b is center shot
library(tidyverse)

tibble() -> simulations
for(simulation_instance in 1:100){

  a <- rnorm(100,0.8,0.05)
  b <- rnorm(100,0.75,0.05)

  sum(a>b)/length(a) -> volume_a
  1-volume_a->volume_b

  mean(a)
  mean(b)

  library(tidyverse)
  pk_data <- tibble(a,b)

  pk_data %>% filter(a>b) %>% summarise(mean_a=mean(a)) %>% pull(mean_a) -> observed_conversion_rate_a
  pk_data %>% filter(b>a) %>% summarise(mean_b=mean(b)) %>% pull(mean_b)-> observed_conversion_rate_b

  simulations <- bind_rows(simulations, tibble(volume_a = volume_a, volume_b=volume_b,
observed_conversion_rate_a=observed_conversion_rate_a,observed_conversion_rate_b=observed_conversion_rate_b ))
}

simulations %>% write_csv("simulation_data.csv")
}

```

From the 200 sample-statistic data-pairs generated, we obtain a *Pearson* correlation coefficient between location-dependent volume and location-dependent conversion-rate of +0.729, and this is significant at extremely small α -levels (e.g., at the $\alpha = 0.000000000000001$ significance level). With very high confidence, the statistical game of chance predicts a positive relationship between location-dependent volume and location-dependent conversion-rate in the empirical data. This theoretical result follows the basketball shot-location distributional results of Ehrlich and Sanders (MIT SSAC, 2024,[17]), who find moderate, significant, positive correlation between location-conditional NBA team-season shot-volume and expected points from 2016-2022.

Let us recap our two main theoretical findings. Under mixed-strategy Nash equilibrium, there is no expected relationship between location-conditional conversion-rate and shot-volume (0 correlation), consistent with the general properties of a *msNe*. Under big-data simulation of a statistical game of chance, we find an equilibrium correlation of +0.729. While these theoretical findings are distinct, we can form a mutually consistent theoretical prediction from the intersection of their complements. Namely:

H₀: If agents are optimizing according to any one or both of these games, we expect a non-negative correlation between location-conditional conversion-rate and shot-volume across all partition subsets.

Under *msNe*, we expect no correlation between volume and conversion-rate for any subset of the partitions. Under the statistical game equilibrium, we expect a positive correlation for all subsets of

the partitions. The intersection of the complement for these theoretical findings (i.e., that which we do not expect) is a negative correlation for some partition subset. We empirically test this theoretical prediction of non-negative correlation against a behavioral explanation. The stacked spatial heat map plots of Figure 3 represent GAM-estimated spatial heat maps of conversion rate and shot volume, respectively, estimated from the sample data at all locations across the goal space. The figure demonstrates a clear spatial contrast between conversion-rate and shot-volume.

Figure 3: Stacked, Spatial Conversion-Rate and Shot-Volume Heat Maps

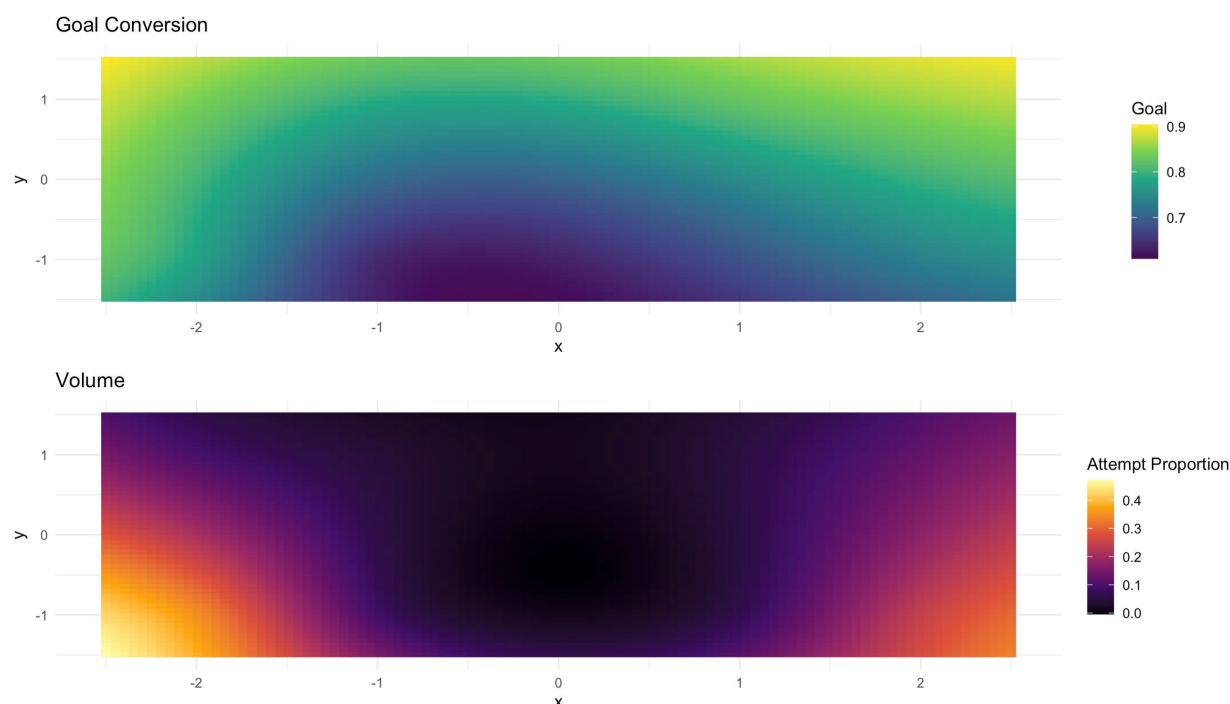
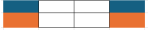


Figure 3 provides fundamental information for our analysis. Moving vertically, conversion-rates rise, along a fixed x-location, from bottom-to-top in the first spatial heat map. However, shot-volumes largely fall moving from bottom-to-top. This contrast is most evident in the shape of the color-based contours within the respective spatial plots of Figure 3. Whereas the contours of the top plot are distinctly concave down, forming an inverted U-shape, the color-based contours of the bottom plot are distinctly concave up, forming a standard U-shape.

Each top partition is unequivocally riskier than its southern neighbor in the bottom-half of the goal. For example, top corner partitions each have 10 feet of post-or-crossbar borders versus 4 feet for bottom-corner partitions. Top-middle partitions each have 6 feet of post-or-crossbar borders versus 0 feet for bottom-middle partitions. Like bumpers in bowling, the ground acts as a guidance mechanism to potentially steer off-target bottom partition shots toward target. There is no such steering mechanism for top partition shots. Hence, it is more challenging in terms of placement to control top shots against borders as compared to bottom shots, *ceteris paribus*, or for a given fixed, partitioned range of x. Further, it is potentially more difficult to control top shots because they simultaneously require both substantial lift and placement, whereas bottom-partition shots depend more specifically on placement. The simultaneous challenge of lift and placement likely puts additional demands on the PK-taker, who is already dealing with increased post-or-crossbar

border-lengths. While a well-placed top-partition shot can decrease the likelihood of a save, it is not a safe shot if a player wishes to appear as a “credible” on-target shooter.

Indeed, we find from the data that the missed goal rate for top-partition PK shots is 12.88% in the data versus 2.48% for bottom partitioned goals. On the other hand, top-partition shots are saved 5.30% of the time versus 23.51% of the time for bottom-partition shots. Top and bottom shots are fundamentally different. Therefore, we subset the partitions to find the correlation between shot-volume and conversion-rate for a) the top and bottom corner partitions and b) the top and bottom middle partitions. That is, $Subset_A = \{Top\ Left, Bottom\ Left, Top\ Right, Bottom\ Right\}$ or  and $Subset_B = \{Top\ Middle\ Left, Bottom\ Middle\ Left, Top\ Middle\ Right, Bottom\ Middle\ Right\}$ or . Under our theoretical findings, no observed correlations should be negative under equilibrium. For $Subset_A$, we obtain a correlation between volume and conversion-rate of (-0.749) . For $Subset_B$, this correlation is -0.997 .

We would like to verify whether these correlations are significantly negative. Unlike the regression data analysis to follow, which is at the shot-level ($n = 536$), this summary correlation analysis is at the partition-level, where there are a total of 8 partitions. Inferential significance testing at such a low sample size is noisy. Using custom resampling code in *R* (see *resampling.R* file in github), therefore, we *bootstrap* resample these correlations 30 times each. Bootstrap resampling is a leading statistical tool for inferential testing in empirical settings that lack either sample size, a counterfactual to the estimate obtained, or both. To do this we input each respective sample correlation coefficient estimate, r , along with its standard error ($\sqrt{\frac{(1-r^2)}{n-2}}$). We then *t*-test the mean for each sampling distribution of resampled correlation values and find that the sampling distribution means are each significantly negative (e.g., at the $\alpha = 0.001$ level). The *t*-statistics for the respective tests are (-7.71) for Partition Subset A and (-10.85) for Partition Subset B. That is, we observe significant, negative correlation from top-to-bottom for both subsets, which provides strong evidence against either the classical or statistical game-theoretic equilibria holding globally, as these require non-negative correlation between any and all partition subsets. Then, there is strong preliminary evidence that behavioral factors are present, possibly with optimizing factors, in determining PK-shot location. Specifically, top shots are more likely to miss the goal-face entirely, and this appears to matter to PK-takers. Our regression and ML regularized Lasso models will further examine whether behavioral and optimizing factors matter to PK-takers conditional on one another. To understand this lack of correspondence, we first run the following shot-level model specifications given in (1) and (2):

$$\text{likelihood shot located to revealed bin}_i = \beta_0 + \beta_1 \cdot \text{expected conversion rate at revealed bin}_i + \beta_2 \cdot \text{expected off-target rate given not-converted at revealed bin}_i + \epsilon_i \quad (5)$$

$$\text{likelihood shot located to revealed bin}_i = \beta_0 + \beta_1 \cdot \text{expected conversion rate at revealed bin}_i + \beta_2 \cdot \text{expected conversion rate squared at revealed bin}_i + \beta_3 \cdot \text{expected off-target rate given not-converted at revealed bin}_i + \epsilon_i \quad (6)$$

The variable *expected off-target rate given not-converted at revealed bin* represents the proportion of other shots to the observed partition that miss the goal given that they were not converted. This variable considers the relative manner in which shots to a given partition are not converted (i.e., proportion of times missed versus saved) and is a behavioral rather than an optimization variable.

It is behavioral, in the sense that the *manner* of not converting a PK-shot does not affect the match score. A miss and a save are equivalent in terms of match-score effect, at zero effect. Then, the materiality of the *manner* of a non-converted PK shot lies squarely in the optics of the shot. The variable *expected conversion rate at revealed bin* considers the conversion-rate for other shots to the observed partition and is an optimization variable in that it affects the match score. The variable *expected conversion rate squared at revealed bin* is the square of the conversion-rate for other shots to the observed partition. This is an optimization variable that allows for a non-linear relationship.

These models are related to a rational expectations model in economics in their interpretation of decision-making. They each take the representative shooter as having an *a priori* rational expectation of the conversion-rate from locating the ball to each respective partition, $x_{1,i}$ in both models, as set by the conversion-rate of all *other* shots in the data located to that partition. The observed shot is left out of the PK-taker's rational expectations forming information set to avoid endogenous outlook and estimation. The shooter also has an *a priori* likelihood of locating the ball to the *ex post* chosen bin (i.e., a likelihood before the choice is made), y_i or the dependent variable in both models, as per rational expectations according to the likelihood that other shots in the data were located to the *ex post* chosen bin. As such, the model considers whether a PK-taker having empirical rational expectations uses partition-dependent conversion-rate expectations to improve or even optimize the PK shot-location distribution across partitions. If the PK-taker behaves thus, there will be a non-negative relationship between expected conversion-rate and shot-frequency across partitions, as is derived and observed in the NBA shot chart analysis of Ehrlich and Sanders (MIT SSAC 2024, [17]). Alternatively, the shooter may consider behavioral factors, such as *a priori expected off-target rate given the shot is not converted at revealed bin i*, where this variable is set with rational expectations as the off-target rate for all other shots to or nearest that partition in the data that were not converted. Table 2 provides the results of estimating the linear (1) and polynomial (2) models using least squares.

Table 2: Linear and Polynomial Regression Model Estimates, Locating PK-Shots

	<i>Dependent variable:</i>	
	likelihood shot located to revealed bin	
	(1)	(2)
expected off-target rate given not-converted at revealed bin	-0.477*** (0.009)	-0.445*** (0.009)
expected conversion rate at revealed bin	1.492*** (0.034)	3.997*** (0.317)
expected conversion rate squared at revealed bin		-1.748*** (0.220)
Constant	-0.794*** (0.024)	-1.687*** (0.115)
Observations	536	536
R ²	0.852	0.868
Adjusted R ²	0.852	0.867
Residual Std. Error	0.047 (df = 533)	0.045 (df = 532)
F Statistic	1,539.231*** (df = 2; 533)	1,166.289*** (df = 3; 532)

Note:

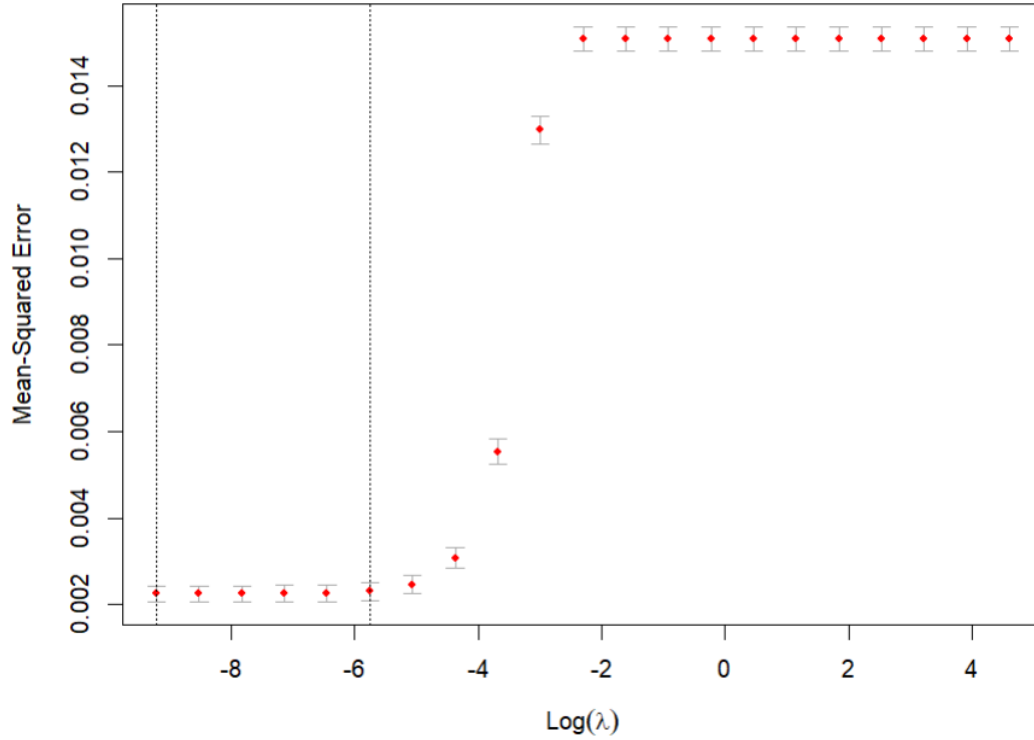
*p<0.1; **p<0.05; ***p<0.01

The models are highly-explanatory: Between 85.2% and 86.8% of the variation in PK-shot location volume is explained by *expected conversion rate at location*, *expected conversion rate squared at location* (Model 2), and *expected off-target rate given not-converted at location*. Each model coefficient is highly significant, suggesting that PK-shot locations are intentional and explainable at the top professional level. This intentionality does not reflect match score optimization alone, however. Both models show that players consider both optimization *and* behavioral factors when choosing a PK-shot location. The representative shooter locates the shot to a partition significantly more as the *expected conversion rate at location* rises but does so significantly less as the *expected off-target rate given not-converted at location* rises.

This suggests that players are not pure optimizers but also consider the optics of missed shots, and this prevents either of the alternative theoretical equilibria from being obtained. This tradeoff between optimization and a behavioral utility for the optics of play also costs their teams expected goals. Model (2) provides equivalent results, both qualitatively and in statistical significance, and also presents evidence that the optimization relationship observed is non-linear. The coefficient on *expected conversion rate at location* attenuates in Model (2) with the inclusion of the quadratic term, but not much. A joint F-test also determines that Model (2) constitutes our best specification. Model (2) also suggests that PK-takers are risk averse, as their partition-dependent shot-volume increases at a decreasing rate in conversion-rate. As the purpose of a PK-attempt is to maximize expected goal (likelihood) directly, the presence of risk-aversion here represents another behavioral factor on the part of the PK-taker. This factor is akin to the seminal findings of [5] and [6] that NFL coaches are behaviorally risk-averse on fourth-down play-calls given that the objective of an American football game, like that of a soccer match, is to maximize point margin.

To robustness check our results against multicollinearity, we also implement an ML-regularized *Lasso* regression estimation for the specification of Model (1) in *R* using the *glmnet* and *selectiveInference* packages. Regularization is a machine learning technique that reduces overfitting in estimation settings that feature multicollinearity. In a Lasso regression, the hyperparameter, λ , balances the tradeoff between estimation bias and variance. At the best-fit hyperparameter, the ML regularized Lasso regression yields coefficients of 1.154 (*expected conversion rate at location*) and -0.387 (*expected off-target rate given not-converted at location*) with respective bootstrapped standard errors of 0.024 and 0.007. Therefore, both coefficients are significant and maintain the same sign as with the linear and polynomial regressions. Figure 4 provides the ML-regularized Lasso model fit at different tuning parameters, λ . This plot shows a low mean-squared error, especially at the best-fit model (0.0022). This indicates that the model provides close predictions of variation in shot-volume by partition given expected conversion-rate and expected miss rate.

Figure 4: ML regularized Lasso model fit at different tuning parameters, λ



Using this ML-technique, we have robustness checked our results against inferential noise from potential non-structural multicollinearity. We will now develop the analysis toward Figure 5, which considers a contour plot of indifference curves derived from the output of the original Model 1. Model (1) yields the following fitted regression model output:

$$\begin{aligned} \text{a priori likelihood shot located to revealed bin}_i &= -0.794 + 1.492 \cdot \\ \text{expected conversion rate at revealed bin}_i &- 0.477 \cdot \text{expected off target rate at revealed bin}_i \end{aligned} \quad (7)$$

Recognizing that the expected off target rate is simply one minus the expected on-target rate, we transform the model as follows, such that we can plot indifference curves on a standard two-good Cartesian plane.

$$\begin{aligned} \text{a priori likelihood shot located to revealed bin}_i &= -0.794 + 1.492 \cdot \\ \text{expected conversion rate at revealed bin}_i &+ 0.477 \cdot \text{expected on target rate at revealed bin}_i \end{aligned} \quad (8)$$

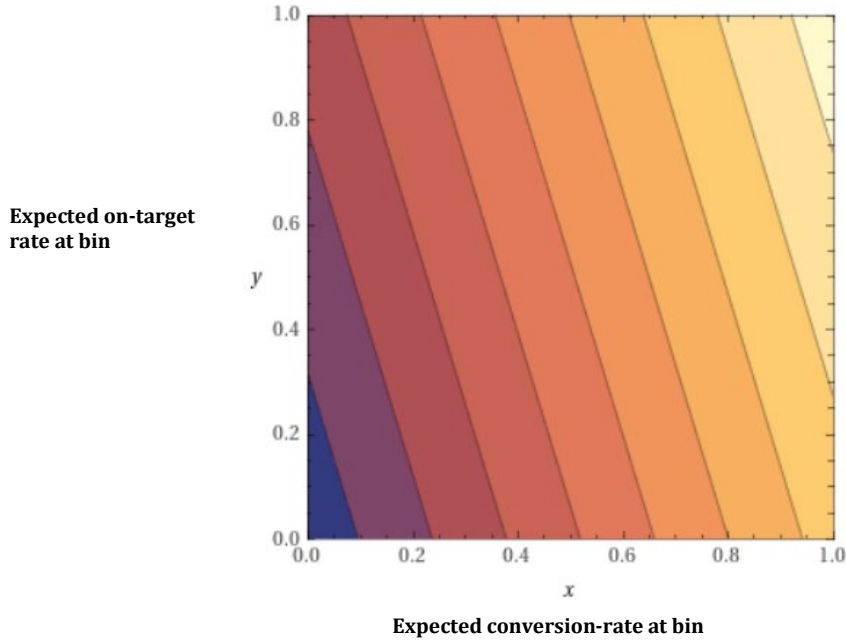
To obtain sets of indifference curves from the latter equation, we simply fix y to some arbitrary value, y_0 , and solve for one x variable in terms of the other (i.e., solve for values of x_1 and x_2 along which y is fixed at y_0).

$$\begin{aligned} y_0 &= -0.794 + 1.492 \cdot \text{expected conversion rate at revealed bin}_i \\ &+ 0.477 \cdot \text{expected on target rate at revealed bin}_i \end{aligned} \quad (9)$$

$$\text{expected on target rate at revealed bin}_i = \frac{y_0 - 0.794 - 1.492 \cdot \text{expected conversion rate at revealed bin}_i}{0.477} \quad (10)$$

We then change y_0 arbitrarily within its possible range to obtain a family of indifference curves rendered in *Mathematica*.

Figure 5: Indifference Curves or Level Sets between *expected conversion-rate at bin* (x-axis) and *expected on-target rate at bin* (y-axis)



This contour plot represents sets of indifference curves for PK-takers (i.e., sets of negatively-sloped (expected conversion-rate, expected on-target rate) tradeoff lines where the shooter is indifferent or equally probable to shoot along a given line). Importantly, these are revealed or regression estimate derived tradeoff lines. That is, they are based on the empirically-revealed shot-location preferences of shooters. The slope of any given indifference curve tells us the tradeoff rate for the representative shooter. The representative shooter would trade or give up a percentage point of conversion rate to raise his or her on-target rate by 3.12 percentage points with indifference according to Model 1. Conversely, the representative shooter would trade or give up a percentage point of on-target rate to increase his or her conversion rate by 0.32 percentage points with indifference according to Model 1. In other words, optimization utility is estimated to be 3.12 times as important to the representative PK-taker on the margin as is behavioral utility.

IV. Conclusion

For penalty kicks in soccer, the classical and statistical game-theoretic models predict a non-negative relationship between goal area partition-conditional shot-volume and conversion-rate for a goal-optimizing penalty kick taker in soccer. However, we find a strong negative relationship between goal area partition subsets of the study data, which considers 536 penalty kicks from the 2015-2020 UEFA Champions and Europa Leagues. The estimated indifference sets and the

underlying inferential statistics of the linear, polynomial, and ML regularized Lasso regression models indicate that penalty-takers (significantly) value both conversion-rate and on-target rate when locating PK-shots. While partial optimizers, PK-takers in soccer deviate from optimal PK-locating strategies in a manner that is consistent with the behavioral valuation of keeping up appearances of highly-skilled play, in this case by limiting the likelihood of missing the goal entirely. Optimization utility is estimated to be 3.12 times as important to the representative PK-taker on the margin as is behavioral utility. The models are highly explanatory, indicating that optimization and behavioral factors explain approximately 85.2% of PK-shot locating variation for top professional players. The negative estimated slope of the indifference sets provides visual evidence that players are revealed to value both conversion-rate and on-target rate. They trade-off between these interests when selecting penalty-shot locations. We conclude strong statistical evidence that suggests penalty-takers represent hybrid decision-makers: part rational-optimizers and part behavioral-agents seeking to keep up appearances of highly-skilled play.

Our polynomial regression also suggests that PK-takers are risk averse, as their partition-dependent shot-volume increases at a decreasing rate in conversion-rate. As the purpose of a PK-attempt is to maximize expected goal (likelihood) directly, the presence of risk-aversion here represents another behavioral factor on the part of the PK-taker. This factor is akin to seminal findings that NFL coaches are behaviorally risk-averse on fourth-down play-calls given that the objective of an American football game, like that of a soccer match, is to maximize point margin.

References

- [1] Gilmore-Jones, J., McRobert, A. P., & Ford, P. R. (2015). A quantitative analysis of penalty kicks in the English Premier League. M. thesis. Research Institute for Sport and Exercise Sciences, Liverpool John Moores University.
- [2] FootyStats. (2024). Most Common Full-Time Scorelines in Football. Available online at: <https://footystats.org/stats/common-score>. Accessed: 08/15/2024.
- [3] TransferMarkt. (2022). Premier League. https://www.transfermarkt.us/premier-league/elfmeterstatistiken/wettbewerb/GB1/saison_id/2022. Accessed 08/11/2024.
- [4] TransferMarkt. (2022). Premier League. https://www.transfermarkt.us/premier-league/elfmeterstatistiken/wettbewerb/GB1/saison_id/2023. Accessed 08/11/2024.
- [5] Romer, D. (2006). Do firms maximize? Evidence from professional football. *Journal of Political Economy*, 114(2), 340-365.
- [6] Goff, B., & Locke, S. L. (2019). Revisiting Romer: Digging deeper into influences on NFL managerial decisions. *Journal of Sports Economics*, 20(5), 671-689.
- [7] Beslic, Stephen (2022). *That's a shot for sisses* – why Shaquille O'Neal never shot underhanded FTs. Available online at: www.basketballnetwork.net/old-school/thats-a-shot-for-sissies-why-shaquille-oneal-never-shot-underhanded-fts. Accessed 08/25/2024.
- [8] Wesley, Colton (2017). Underhanded free throws work, so why don't players shoot them? Available online at: www.detroitjockcity.com/2017/05/31/underhanded-free-throws-work-dont-players-shoot. Accessed 08/25/2024.

- [9] Pelechrinis, K. (2019). Winning in Basketball with Data and Machine Learning. Syracuse University Sport Analytics Research Seminar. November 15, 2019.
- [10] Almeida, C. H., Volossovitch, A., & Duarte, R. (2016). Penalty kick outcomes in UEFA club competitions (2010-2015): the roles of situational, individual and performance factors. *International Journal of Performance Analysis in Sport*, 16(2), 508-522.
- [11] Osborne, M.J., Rubinstein, A.: A Course in Game Theory. MIT Press 1994.
- [12] Nash Jr, J. F. (1950). Equilibrium points in n-person games. *Proceedings of the national academy of sciences*, 36(1), 48-49.
- [13] S. Kakutani, "A generalization of Brouwer's fixed point theorem," Duke Math. J., vol. 8, no. 3, pp. 457-459, 1941.
- [14] Muth, J. F. (1961). Rational expectations and the theory of price movements. *Econometrica: journal of the Econometric Society*, 315-335.
- [15] Lucas Jr, R. E., & Prescott, E. C. (1981). Optimal investment with rational expectations. *Rational expectations and econometric practice*, 1, 127-156.
- [16] Sargent, T. J. (1972). Rational expectations and the term structure of interest rates. *Journal of Money, Credit and Banking*, 4(1), 74-97.
- [17] Ehrlich, J. and Sanders, S. (2024). Estimating NBA Team Shot Selection Efficiency from Aggregations of True, Continuous Shot Charts: A Generalized Additive Model Approach. Available online at: www.sloansportsconference.com/research-papers/estimating-nba-team-shot-selection-efficiency-from-aggregations-of-true-continuous-shot-charts-a-generalized-additive-model-approach. Accessed 08/30/2024.