LUCK AND THE LAW: QUANTIFYING CHANCE IN FANTASY SPORTS AND OTHER CONTESTS*

DANIEL GETTY†, HAO LI‡, MASAYUKI YANO§, CHARLES GAO¶, AND A. E. HOSOI‖

Abstract. Fantasy sports have experienced a surge in popularity in the past decade. One of the consequences of this recent rapid growth is increased scrutiny surrounding the legal aspects of the game which typically hinge on the relative roles of skill and chance in the outcome of a competition. While there are many ethical and legal arguments that enter into the debate, the answer to the skill versus chance question is grounded in mathematics. Motivated by this ongoing dialogue we analyze data from daily fantasy competitions played on FanDuel during the 2013 and 2014 seasons and propose a new metric to quantify the relative roles of skill and chance in games and other activities. This metric is applied to FanDuel data and to simulated seasons that are generated using Monte Carlo methods; results from real and simulated data are compared to an analytic approximation which estimates the impact of skill in contests in which players participate in a large number of games. We then apply this metric to professional sports, fantasy sports, cyclocross racing, coin flipping, and mutual fund data to determine the relative placement of all of these activities on a skill-luck spectrum.

Key words. Fantasy sports, chance, statistics of games

AMS subject classifications. 62-07, 62P99, 91F99

1. Introduction.

In his 1897 essay, Supreme Court Justice Oliver Wendell Holmes famously wrote: “For the rational study of the law the blackletter man may be the man of the present, but the man of the future is the man of statistics ...” [11]. This view has proven to be prescient as recent trends show that legal arguments grounded in data analysis are becoming increasingly common [7]. In light of this shift – although historical and ethical arguments remain the purview of legal professionals – physicists, mathematicians, and others well versed in data science have an obligation to provide rigorous mathematical foundations to ground these statistical legal debates.

One such debate that is currently being argued in the courts involves fantasy sports, a game in which participants assemble virtual teams of athletes and compete based on the athletes’ real-world statistical performance. Fantasy sports have experienced a surge in popularity in the past decade. The Fantasy Sports Trade Association [2] estimates that 56.8 million people played fantasy sports in 2015 (up from 41.5 million in 2014) and that the concomitant economic impact of the industry is on the order of billions of dollars per year. One of the consequences of the recent rapid growth in activity is an increased scrutiny regarding the legal aspects of the game. In particular, contests that involve online exchange of funds are now subject to the Unlawful Internet Gambling Enforcement Act of 2006 (UIGEA) which regulates online financial transactions associated with betting or wagering [3]. Currently, the

*Submitted to the editors September 25, 2016.

Funding: This work was partially funded by FanDuel. FanDuel did not exercise any editorial control over this paper.

†Department of Aeronautics and Astronautics, Massachusetts Institute of Technology, Cambridge, MA 01239.

‡Department of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139.

§University of Toronto Institute for Aerospace Studies, Toronto, ON M3H 5T6, Canada.

¶Sloan School of Management, Massachusetts Institute of Technology, Cambridge, MA 02139.

‖Department of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139 (peko@mit.edu)

UIGEA excludes fantasy sports stating that the definition of a "bet or wager" does *not* include "participation in any fantasy or simulation sports game" [3]. However, at the time of this writing eight states do not allow fantasy sports competitions for cash.

In general, legal questions surrounding classification of games as "bets" or "wagers" hinge on whether the outcome of the game is determined predominantly by skill or by chance. Typically "skill" is defined as the extent to which the outcome of a game is influenced by the actions or traits of individual players, compared to the extent to which the outcome depends on random elements. Note that here we are considering the role of skill in the game framework, i.e. our goal is to quantify the utility of a players' abilities in general, rather than to measure the skill of individual players.

In a fantasy sports contest, players manage teams that accrue points based on the statistics of real athletes. For example, a fantasy football player with receiver X on their roster earns points every time X makes a catch in a professional football game. Rules for constructing a fantasy team roster – and hence strategies for assembling optimal line-ups – vary by league. In this study, we analyze data from *salary cap* games in which each player has a fictional dollar amount they can spend, and athlete "salary" values are set by the game provider.[1]

While there have been relatively few empirically-based investigations on the roles of skill and chance in fantasy sports, studies exist to determine whether skill is a distinguishing factor among NFL kickers [19], whether the outcome of shootouts in hockey are primarily determined by luck [12], and whether perceived "streaks" in basketball should be attributed to chance or to "the hot hand" [25, 27, 28]. In addition, many authors have explored whether scoring patterns in basketball [4, 9, 24], baseball [17, 24], cricket [23], soccer [10], tennis [13, 22] and Australian rules football [14] display statistical signatures of random processes. Perhaps the most relevant, extensive (and hotly debated) body of work is the analysis of the relative roles of skill and chance in poker [6, 5, 26, and references therein]. Unfortunately, most of these analyses cannot be applied directly to fantasy sports which differ from card games in at least one critical aspect: in card games, it is a relatively straightforward combinatorial exercise to estimate the probability of every possible outcome (e.g. how likely is a flush). This is not the case for fantasy sports in which performance on any given day is coupled to a host of factors such as weather, skill of opponent, injury, home versus away games, etc. Owing to this added layer of complexity, it is unclear whether the bulk of previous poker analyses can be adapted to fantasy games.

Instead, we take a data-driven approach analogous to the one proposed by Levitt *et al.* [16, 15]. Their study was performed in the context of poker but, unlike the combinatorial approaches of e.g. [6], it does not require prior knowledge of outcome probabilities and hence can be easily adapted to other activities that combine skill and chance. Our strategy will be to examine data from the online fantasy sports platform FanDuel, and test whether statistical outcomes are consistent with expected outcomes associated with games of chance. If the measured outcomes deviate significantly from those we expect in a contest of pure chance, we can quantify the extent of the deviation to place fantasy sports and other activities on a skill-luck spectrum.

Levitt *et al.* [16] note that there are at least four tests that can be applied to distinguish games of pure chance from those involving skill. The authors propose that tests can be framed as the following four questions:

1. Do players have different expected payoffs when playing the game?

[1] Here, and throughout this paper, we use the term *player* to refer to a person participating in a fantasy sports competition whereas *athlete* refers to professional athletes.

MLB	P	C	1B	2B	3B	SS	OF	OF	OF
NBA	PG	PG	SG	SG	SF	SF	PF	PF	C
NFL	QB	RB	RB	WR	WR	WR	TE	K	D
NHL	LW	LW	RW	RW	C	C	D	D	G

TABLE 1

FanDuel 2013/2014 roster rules. MLB: P = pitcher, C = catcher, 1B = 1st baseman, 2B = 2nd baseman, 3B = 3rd baseman, SS = shortstop, OF = outfielder. NBA: PG = point guard, SG = shooting guard, SF = small forward, PF = power forward, C = center. NFL: QB = quarterback, RB = runningback, WR = wide receiver, TE = tight end, K = kicker, D = defense. NHL: LW = left wing, RW = right wing, C = center, D = defenseman, G = goalie.

2. Do there exist predetermined observable characteristics about the player that help one to predict payoffs across players?
3. Do actions that a player takes in the game have statistically significant impacts on the payoffs that are achieved?
4. Are player returns correlated over time, implying persistence in skill?

If the answer to *all four* questions is "no," then we can conclude that the game under consideration is a game of chance. One of the many appealing aspects of this test is that the analysis can be framed in terms of inputs (player actions) and outputs (win-loss records), hence the game itself can be treated as a black box and the relative roles of skill and luck can be quantified irrespective of the detailed rules of the game.

2. Empirical tests of skill applied to fantasy sports data. To estimate the relative roles of skill and chance in these contests, we analyze data from FanDuel, currently one of the largest providers of daily fantasy sports. We consider two types of daily fantasy games − head-to-head (H2H) and 50/50 competitions − associated with four sports leagues − Major League Baseball (MLB), the National Basketball Association (NBA), the National Football League (NFL), and the National Hockey League (NHL). In H2H competitions, the player pits his team against a single opponent; both players pay the same amount to play and the winner takes all (minus the overhead to the host site). In a 50/50 league a pool of players each pay the same entry fee to enter the competition; the top half of scorers in the fantasy league each receive the same payout (roughly double what they put in), while the bottom half receives nothing. When a player elects to play in a *game* (e.g. a particular 50/50 competition), they may pay to submit multiple *entries*. Hence each player has a win fraction associated with each game defined as their fraction of winning entries. These entries are typically not independent and players may submit multiple copies of the same entry.

FanDuel provided us with 12 sets of data [1]. The first four were anonymized results from H2H competitions for MLB, NBA, NFL, and NHL. Each entry in the data set represented the performance of one user in one game (G_i) and contained: a user ID (UID), number of entries submitted by UID in G_i, number of winning entries for UID in G_i, average score (averaged across all UID's entries in G_i), and the top score for UID in G_i. The next four data sets contained similar information for 50/50 competitions. The final four datasets contained athlete performance data. Each entry in the dataset included: athlete name, date of competition, team, FanDuel "salary" on the date of that particular competition, position (summarized in Table 1), and number of FanDuel points scored by that athlete in that particular competition.

	MLB	NBA	NFL	NHL
# plr (H2H)	89,338	107,796	190,562	20,802
# plr (50/50)	130,515	158,532	312,263	30,607
Max games/plr	686	383	155	271
Max entries/plr	60,200	64,287	49,657	21,209
Salary cap	35,000	60,000	60,000	55,000
# athletes	1,261	501	584	891

TABLE 2
Overview of the FanDuel dataset. Note that players (plrs) may have multiple entries for each game.

We considered two full seasons (2013/14 and 2014/15) of all FanDuel H2H and 50/50 contests for NBA, NHL, NFL, and MLB. Each fantasy team for all sports contains nine athletes (see Table 1). The salary caps, number of fantasy players, and number of athletes in our dataset are summarized in Table 2. In the following we use this data to test questions 1, 3, and 4 proposed by Levitt *et al.* In all cases we found no significant differences between the H2H and 50/50 data so the results presented herein were obtained using the combined datasets unless otherwise noted. Since our data is anonymized we do not have information on predetermined player characteristics; hence we do not address question 2.

2.1. Expected payoff. In a game of chance, the expected payoff for all players is the same. To test whether this is true of our data, we divide each of our datasets – fantasy NBA, MLB, NFL, and NHL – into five subsets according to the number of entries N_i played by the ith player. The first group contains players who have submitted the fewest number of entries and the fifth group contains players who have submitted the largest number of entries. In the FanDuel playing population, we observe that the number of players, m, who have played n_i games decays approximately exponentially (i.e. most players play only a few games, see Figure 4), hence ranges were selected to reflect a logarithmic distribution such that the first group contains 90% of the players, the second group 90-99%, the third 99-99.9%, etc. If the measured win fraction distribution varies across these five subsets in a statistically significant manner we can conclude that something other than chance played a role in the outcome of the contest. Here the win fraction of the ith player is computed as

$$w_i = \frac{1}{n_i} \sum_{j=1}^{n_i} x_{ij} \tag{1}$$

where x_{ij} represents the fraction of winning entries in the j-th game: $0 \leq x_{ij} \leq 1$ and $x_{ij} = 1$ if all of player i's entries in the j-th game were winners.

Some care must be taken in this analysis as it can be argued that players who win their initial games (whether by skill or by luck) may be more likely to keep playing; conversely, players on a losing streak may be more likely to quit. To determine whether this is the case, we computed the average win fraction across the population for the n_ith game (i.e. the last game the player played before they quit), the $n_i - 1$ game, (i.e the second-to-last game the player player before they quit), etc. These data are summarized in Figure 1 (Left) which shows a "quitting boundary layer" indicating that players are indeed more likely to quit after a string of losses. In our dataset,

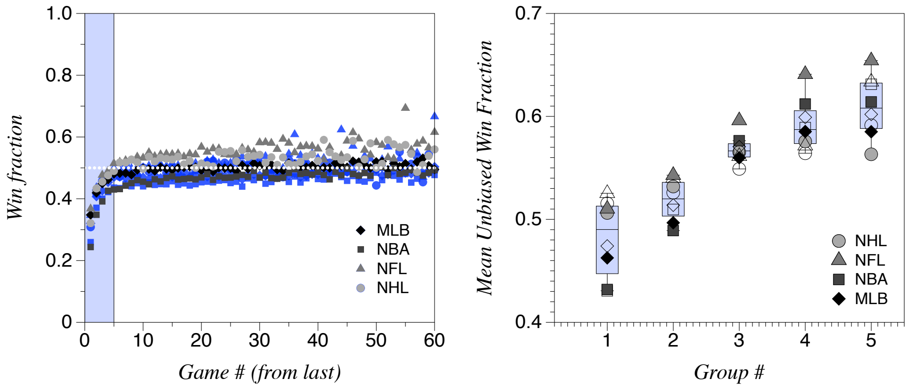

Fig. 1. *(Left) Win fraction associated with each player's last game (game # = 1), second to last game (game # = 2), etc. averaged across the entire playing population. Grey symbols correspond to 50/50 games, blue symbols to H2H games from the 2013/14 season; shapes correspond to different sports as summarized in the legend. The blue shaded region on the left corresponds to the "quitting boundary layer." (Right) Mean win fraction for fantasy MLB, NBA, NFL, and NHL in which the boundary layer points have been removed for each player. Group 1 corresponds to players who have played the fewest number of games; group 5 corresponds to players who have played the largest number of games. Filled symbols correspond the 2014/15 season; empty symbols correspond to the 2013/14 season. Light blue boxes are a standard IQR box-and-whisker plots.*

the boundary layer, n_{BL}, is measured to be approximately five games wide in all four sports. This introduces a bias in the calculation since the boundary layer accounts for a larger fraction of games in the first group in which the population has played the fewest number of games. To correct for this bias, we remove the last five games played by every player and compute the win fraction of the remaining games:

$$w_{i,unbiased} = \frac{1}{n_i - n_{BL}} \sum_{j=1}^{n_i - n_{BL}} x_{ij}. \tag{2}$$

The removal of the boundary layer increases the average win fraction significantly in the first group but has only a minor impact on groups 2–5.

A number of trends are clearly observable in the data summarized in Figure 1 (Right). First, players who play the fewest games systematically underperform the other groups with an average win fraction of 0.48 (averaged over all four sports). In contrast, in the cohort that played the most games, the mean win fraction increased to 0.61 (averaged over all four sports). Thus these data are not consistent with a game of pure chance. (Additional details are available in Table 4 in the Appendix.)

2.2. Effect of player action. To test the effects of player actions, we compare outcomes of real players with those of a league of players who draw randomly from all possible line ups. Ideally we would like to compare the distribution of scores generated by real fantasy players with the distribution of scores for all possible line-ups, however the combinatorics associated with generating all possible line-ups proved to be computationally intractable. Instead we estimate the "all possible line-up" distribution using a Monte Carlo approach. Two strategies were tested to construct the Monte Carlo rosters. In the first, athletes for each roster slot were picked randomly from a normal distribution of salaries, centered at one ninth of the salary cap. In the second, the center of the salary distributions varied by position; for example, if on average quarterbacks cost twice as much as kickers, then the mean of the normal salary distribution for quarterbacks was set to be twice as much as the mean of the normal salary distribution for kickers, with the constraint that the sum of the means across all positions must be equal to the salary cap.

One hundred random line-ups that conform to the rules of the game were generated for each day of competition. In all cases including football, only daily competitions were considered which corresponds to a range of 51 (NFL) – 182 (MLB) competitive days per year depending on the sport. Each roster is checked to see if the total salary is below the salary cap and above a minimum threshold. In the data shown here, the threshold is set to 85% of the salary cap which roughly maximizes the mean score of the random line-ups using FanDuel data. (If the threshold is too low, cheap line-ups can skew the distribution towards low scoring rosters; if the threshold is too high, the constraint is too rigid and high performing rosters are missed.) Randomly generated rosters that satisfy these constraints are accepted, all others are rejected. The process is repeated until 100 acceptable rosters per day were generated. We found very little difference between the two Monte Carlo strategies, i.e. equal distribution versus position-weighted distribution. Results for the weighted distribution are shown in Figure 2.

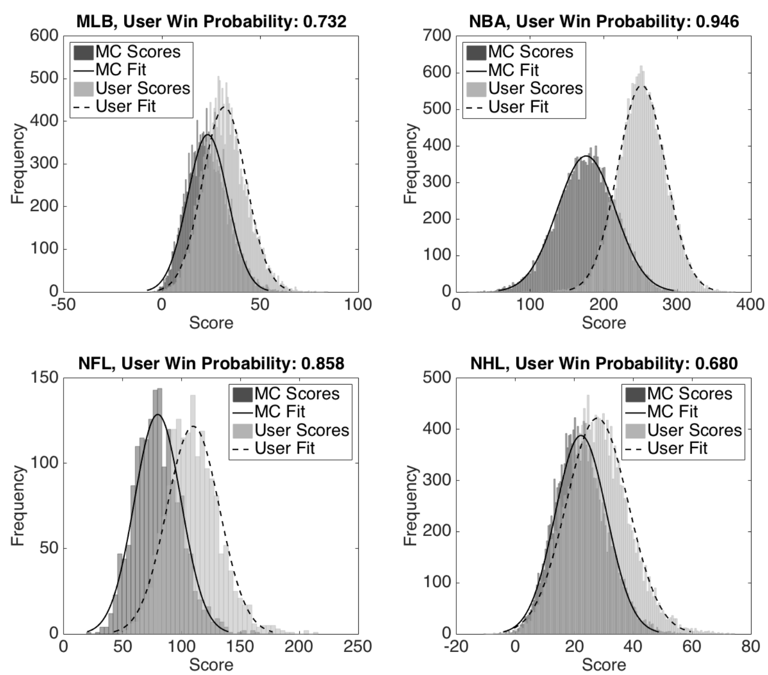

FIG. 2. *Comparison of scores from line-ups constructed by real FanDuel players (light grey) with line-ups from the Monte Carlo simulation (dark grey). Resulting distributions are fit with a normal distribution to estimate confidence intervals on user win probabilities.*

In all cases FanDuel players beat the Monte Carlo simulation with user win probabilities ranging from 62% (NHL, equal distribution) to 95% (NBA, weighted distribution) as summarized in Table 3 suggesting that player actions do indeed influence the outcome of the game.

	MLB	NBA	NFL	NHL
User win prob (weighted)	0.732	0.946	0.864	0.680
95% CI (weighted)	±0.007	±0.004	±0.016	±0.007
User win prob (equal)	0.781	0.898	0.809	0.618
95% CI (equal)	±0.006	±0.005	±0.019	±0.007

Table 3
Summary of win probability of user generated rosters versus Monte Carlo generated rosters.

2.3. Persistence. To address the question of persistence, we begin with the hypothesis that skill is an intrinsic quality of a fantasy player and does not change significantly over the course of the season. If this is the case we expect to observe a distribution of underlying skill across the playing population in which the win fraction of each individual player in the first half of the season is correlated with that player's win fraction in the second half. To determine whether this is consistent for FanDuel players, we plot the win fraction for the first half of the season versus the win fraction for the second half of the season for each player. (Note that here and in all subsequent calculations, the quitting boundary layer has been removed.) These data are shown in the scatter plots in Figure 4; each circle represents one FanDuel player and the size of the point represents the number of games played by that particular player.

In order to quantify the role of skill in determining the outcomes of the competitions represented in these plots, we seek a metric with the following properties. Ideally the accuracy of the metric should improve as number of contests per player grows, $n_i \to \infty$, and as the number of players grows, $m \to \infty$. In particular, the metric must capture the expected behaviors at the two extremes: competitions of pure luck and competition of pure skill. In contests of pure luck, the expected outcome of every player is the same. Hence – assuming a zero-sum game in which players are playing against one another – as the number of games per player increases, win fractions for the first half of the season versus win fractions for the second half of the season for all players converge to a single point at $(1/2, 1/2)$. Conversely, in a skill dominated competition, we expect to observe a distribution of skill across the playing population; in this case we expect the data converge to a line with slope one as the number of games per player increases.

Sketches of expected user win fractions for the first half of the season versus user win fractions for the second half of the season for high and low values of n_i and m are shown in Figure 3. To characterize these distributions, perhaps the most obvious metrics to try are the Pearson product-moment correlation coefficient or a standard linear regression. However, both of these are problematic. The linear regression fails to accurately capture the extreme of pure luck since, in the limit that the number of games per player becomes large, all of the data collapses onto a single point. The Pearson product-moment correlation coefficient is problematic as it cannot distinguish between lines of different slopes. In our case, in the limit of pure skill, we expect a line of slope 1, whereas lines with different slopes, no matter how well correlated, are not representative of intrinsic skill. Hence we seek an alternate measure. Motivated by the expected outcomes shown in Figure 3, we propose to use the ratio of the variance along the diagonal in our scatter plots (denoted as S in Figure 3), to the variance in the orthogonal direction (denoted as T in Figure 3) as a measure of the relative importance of skill and luck in a competition.

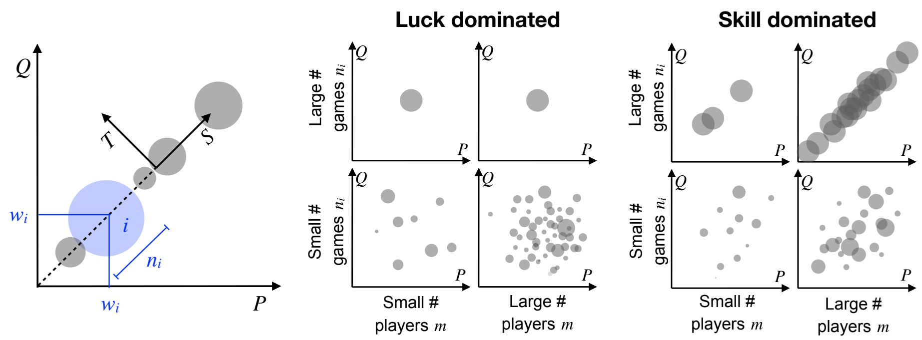

FIG. 3. *(Left) Sketch of expected outcomes for first half win fractions P (horizontal axes) versus second half win fractions Q (vertical axes) along with rotated coordinate system (S, T). (Center and Right) Sketches of expected distributions for games dominated by luck and games dominated by skill in the limit of small and large numbers of contests per player (n_i) and small and large numbers of players (m). In each sketch, points represent individual FanDuel players in which the size of the point represents the number of games played by the ith player, n_i.*

To evaluate this quantity, we characterize each player by two numbers: n_i, the number of games played by player i, and w_i the win fraction of player i as computed in Equation (1). Note that in the following analysis we will be computing quantities associated with a distribution of m players where the ith player constitutes a distribution of n_i games. Hence quantities with the subscript i refer to individual players while quantities with no subscript refer to aggregates over all players. Next, we introduce random variables P_i and Q_i associated with the win fraction of player i in the first and second half of the season, respectively. If the game is truly random, then P_i and Q_i have the same distribution and $E[P_i] = E[Q_i] = 1/2$ (where $E[X]$ indicates the expected value of X); however, if the outcome of the game is primarily determined by skill, then $E[P_i] = E[Q_i] = w_i$, where w_i represents the true underlying skill of player i.

Rotating the (P, Q) coordinate system by $\pi/4$ and shifting the origin, we introduce the transformed random variables as sketched in Figure 3:

$$S_i = \frac{1}{\sqrt{2}}(P_i + Q_i - 1) \tag{3}$$

$$T_i = \frac{1}{\sqrt{2}}(Q_i - P_i). \tag{4}$$

Here T_i represents the difference in the win fraction distribution between the first and second halves of the season, and S_i represents the variation of P_i and Q_i from the nominal value of $1/2$. Note that, if the game is truly random, then in the new coordinate system $E[S_i] = E[T_i] = 0$; however, if the outcome of the game is determined by skill, then $E[T_i] = 0$ and $E[S_i] = (2w_i - 1)/\sqrt{2}$, where w_i varies according to the skill level of the individual player. In games that combine skill and chance, we expect the measured values to lie somewhere between these two extremes.

To characterize the role of skill in determining the outcome of the game, we compute a weighted variance in S and T over the aggregate of all players

$$A \equiv \sigma_S^2 = \frac{1}{m}\sum_{i=1}^{m}(S_i - E[S])^2\theta_{S_i} \tag{5}$$

$$B \equiv \sigma_T^2 = \frac{1}{m}\sum_{i=1}^{m}(T_i - E[T])^2\theta_{T_i}, \tag{6}$$

where $E[S] = (1/m)\sum_i^m S_i = 0$, $E[T] = (1/m)\sum_i^m T_i = 0$, and θ_{S_i} and θ_{T_i} are weighting functions that reflect the our confidence in the ith data point. Finally, we define the quantity

$$R^* = 1 - \frac{B}{A} \tag{7}$$

which provides a single metric to quantify the relative role of skill and chance in determining the outcome of the game. For games that are truly random, $E[R^*] = 0$; for games that are purely skill-based, $E[R^*] = 1$. [2]

To estimate the distribution, expected value, and variance of R^* from real data, we first compute the sample mean estimate for each player associated with the win fraction in the first and second half of the season, respectively, using

$$\hat{p}_i = \frac{1}{n_i/2}\sum_{j=1}^{n_i/2} x_{ij}, \qquad \hat{q}_i = \frac{1}{n_i/2}\sum_{j=n_i/2+1}^{n_i} x_{ij}. \tag{8}$$

Next we model the skill of each player, as reflected by the win fraction, by a normal distribution. Towards this end, we only consider players that satisfy the condition under which the binomial distribution may be well-approximated by a normal distribution namely

$$n_i w_i/2 > 5 \quad \text{and} \quad n_i(1 - w_i)/2 > 5. \tag{9}$$

In this case, the variance in the half season win fraction of the ith player can be approximated by

$$\sigma_i^2 = \frac{2w_i(1 - w_i)}{n_i}. \tag{10}$$

Note that if skill is a persistent quality that is intrinsic to the player, then the win fraction in the first and second half of the season should be equal – namely $\hat{p}_i$ and $\hat{q}_i$ should both approach w_i as the number of games per player becomes large – and players can be represented by points along the diagonal as sketched in Figure 3 (Left).

We can now directly compute the R^* value associated with FanDuel data. At this point we are not computing a distribution, rather the particular instance of R^* that was observed in the 2013/14 and 2014/15 seasons. Rotating $\hat{p}_i$ and $\hat{q}_i$ as defined above to shift to $\hat{S}_i$ and $\hat{T}_i$ coordinates, weighting the ith data point by the variance $\theta_{S_i} = \theta_{T_i} = 1/\sigma_i^2$, and using $E[S] = E[T] = 0$, we compute

$$\hat{A} = \frac{1}{m}\sum_{i=1}^{m}\frac{\hat{S}_i^2}{\sigma_i^2} \qquad \hat{B} = \frac{1}{m}\sum_{i=1}^{m}\frac{\hat{T}_i^2}{\sigma_i^2} \tag{11}$$

[2] Here we choose $R^* = 1 - B/A$ as opposed to $R^* = B/A$ as game designers found it more intuitive to associate games of skill with $R^* = 1$.

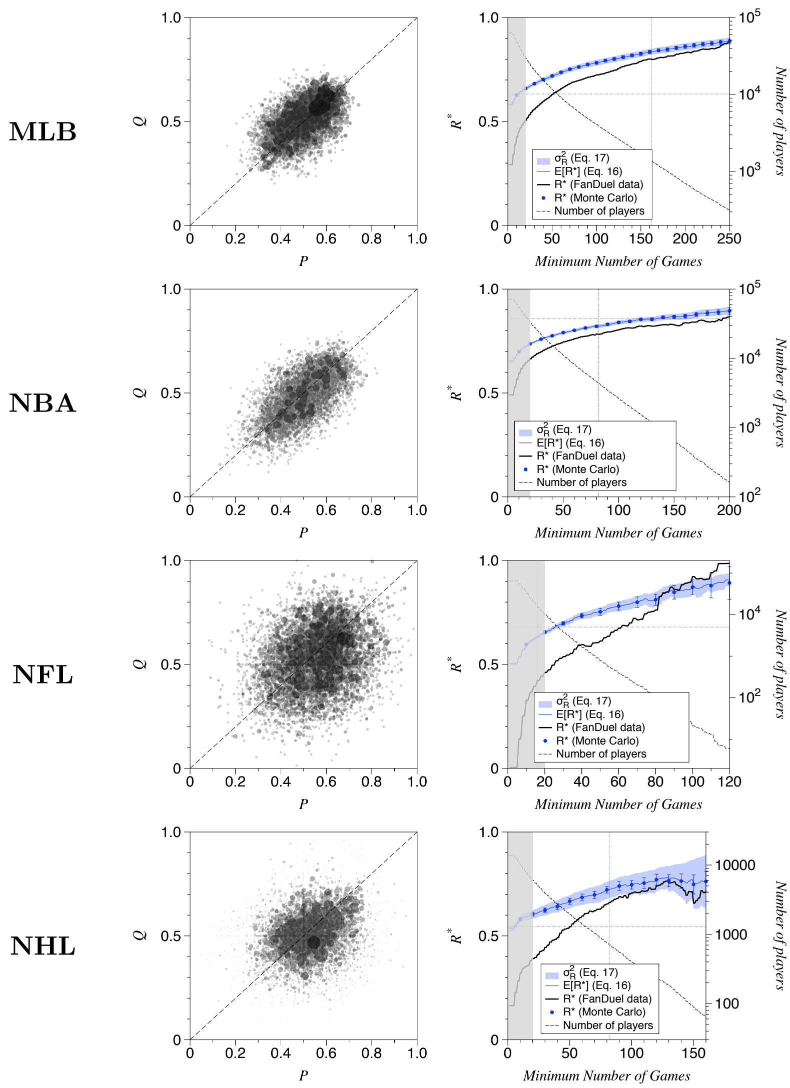

FIG. 4. *Left: Scatter plot of P versus Q for FanDuel players who have played a minimum of G_{min} games. To improve readability in the plots, G_{min} was selected for each sport to display approximately 5000 data points; $G_{min,MLB} = 85$, $G_{min,NBA} = 75$, $G_{min,NFL} = 25$, $G_{min,NHL} = 20$. Each circle represents a single FanDual player; size of the circle represents the number of games played. Right: R^* value calculated from FanDuel data (solid black line), expected value of R^* from Monte Carlo simulations (blue filled circles; error bars represent standard deviation across Monte Carlo trials), computed error (blue shaded region), and number of players in the FanDuel population (dashed line). Vertical dotted line represents the number of games in the corresponding professional sports season and the horizontal dotted line represents the R^* value calculated for the 2010–2015 seasons corresponding to the relevant professional league.*

and find $\hat{R}^* = 1 - \hat{B}/\hat{A}$. This observed value of $\hat{R}^*$ for the 2013/14 and 2014/15 NBA, MLB, NFL, and NHL FanDuel players is shown by the solid black lines in Figure 4. In each plot we consider a range of playing populations defined by the minimum number of games per player represented along the horizontal axis (e.g. if the minimum number of games is 100 then we discard players who have played 99 games or less). The dashed line (right vertical axis) represents the number of players in each population which is surprisingly well-approximated by an exponential distribution in all four sports suggesting that the measured value of R^* is dominated by players who have played the fewest number of games in the sample.

While this calculation serves as a starting point to estimate the role of skill in a particular FanDuel season, ideally we would like to estimate the expected value of R^* (and the variance, etc.) if a large number of seasons were played with the same player population. Towards this end, we model the win fraction of each player in the first and second half by random variables $P_i \sim \mathcal{N}(w_i, \sigma_i^2)$ and $Q_i \sim \mathcal{N}(w_i, \sigma_i^2)$ where $\mathcal{N}(\mu, \sigma^2)$ indicates a normal distribution with mean μ and variance σ^2. Applying the rotated coordinate system transform to obtain the estimated win fractions as before we find

$$S_i = \frac{1}{\sqrt{2}}(P_i + Q_i - 1) \sim \mathcal{N}(\mu_{S_i}, \sigma_i^2) \tag{12}$$

$$T_i = \frac{1}{\sqrt{2}}(Q_i - P_i) \sim \mathcal{N}(\mu_{T_i}, \sigma_i^2) \tag{13}$$

where S_i and T_i are also normally distributed random variables with means given by $\mu_{S_i} = (2w_i - 1)/\sqrt{2}$ and $\mu_{T_i} = 0$. The quantities $\hat{A}$ and $\hat{B}$ can now be computed as in Equation (11) with $\hat{S}_i$ and $\hat{T}_i$ drawn from the specified normal distribution.

In practice, we may estimate the distribution of R^* using a Monte Carlo approach. For each player we can generate a "season" (i.e. x_{ij}'s) with the constraint that the simulated season must have the same n_i and w_i as the real player's season. From this data we can compute a specific instance of $\hat{R}^*$. This process is repeated to construct a distribution which can then be used to estimate quantities of interest, such as the mean, confidence region, etc. Results from the Monte Carlo simulations for NBA, MLB, NFL, and NHL FanDuel players are shown in blue filled symbols in Figure 4. The blue dots represent the mean value of R^* computed for 100 Monte Carlo seasons. Error bars on the symbols represent the standard deviation across the 100 trials.

We may also estimate both the expected value and variance of R^* directly from the player data by approximating $E[R^*] = E\left[1 - B/A\right] \approx 1 - E[B]/E[A]$. Because $S_i/\sigma_i \sim \mathcal{N}(\mu_{S_i}/\sigma_i, 1)$, the random variable A is distributed as a noncentral chi-squared distribution with the parameters m and $\lambda_S = \sum_{i=1}^m (\mu_{S_i}/\sigma_i)^2$. Similarly, the random variable B is distributed as a noncentral chi-squared distribution with the parameters m and $\lambda_T = \sum_{i=1}^m (\mu_{T_i}/\sigma_i)^2$. Hence

$$E[A] = \frac{1}{m}\left(m + \lambda_S\right) = 1 + \frac{1}{m}\sum_{i=1}^{m} \frac{(2w_i - 1)^2 n_i}{2w_i(1 - w_i)} \tag{14}$$

$$E[B] = \frac{1}{m}\left(m + \lambda_T\right) = 1, \tag{15}$$

from which $E[R^*]$ can be computed as

$$E[R^*] = 1 - \left[1 + \frac{1}{m}\sum_{i=1}^{m} \frac{(2w_i - 1)^2 n_i}{4w_i(1 - w_i)}\right]^{-1}. \tag{16}$$

The expected value of R^* computed directly from Equation (16) using FanDuel player data is shown in Figure 4 as solid blue lines.

Similarly, the variance can be estimated directly from player data by propagating the uncertainty associated with the fact that each player only plays a finite number of games:

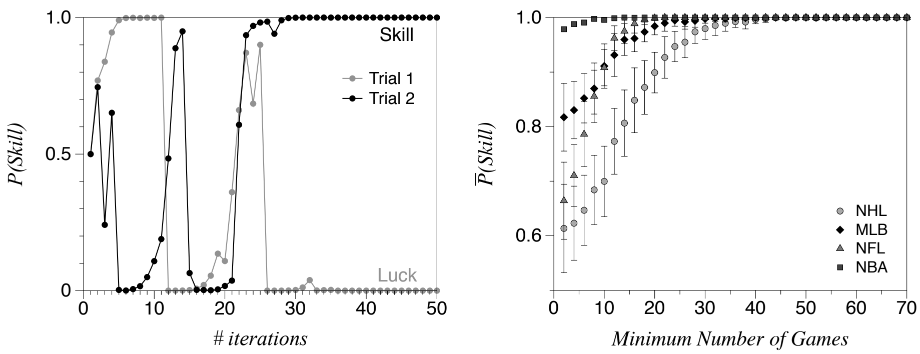

FIG. 5. *(Left) Sample evolution of* $P(\text{Skill})$ *as more player data are added to update the prior for players who have played at least 10 fantasy MLB games. (Right) The probability that the outcome of the contest is determined by skill,* $P(Skill)$ *as a function of the minimum number of games in the playing population. Data is shown for all four sports MLB, NBA, NFL, and NHL.*

$$\sigma_R^2 \approx \sum_{i=1}^{m} \left(\frac{\partial \hat{R}^*}{\partial \hat{w}_i} \right)^2 \sigma_i^2 = \frac{m^2}{8} \left(m + \lambda_S\right)^{-4} \sum_{i=1}^{m} \frac{n_i(2\hat{w}_i^2 - 1)^2}{\hat{w}_i^3(1-\hat{w}_i)^3}. \tag{17}$$

This estimate is indicated by the shaded light blue regions in Figure 4. In all four sports, both the analytic estimate and the Monte Carlo simulation provide a reasonable approximation to the FanDuel data for players that have played more than approximately 100 games. For players with fewer games, Equation (16) and the Monte Carlo simulations over-predict the measured value of R^* suggesting that the assumption that players are well-represented by a normal distribution with mean w_i and variance σ_i breaks down for small n_i.

3. A Bayesian approach. As a final consistency check, we analyze the data from a Bayesian perspective. First, we associate with any given competition a random variable, which characterizes the probability that the competition is a game of luck or a game of skill, with a sample space $\{\text{Luck}, \text{Skill}\}$. Our goal is to infer the probability that the competition is a game of luck (or skill) based on observed data. To this end, we appeal to Bayes' theorem:

$$P(\text{ Skill } | w_i) = \frac{P(w_i | \text{ Skill }) P(\text{ Skill })}{P(w_i)} \tag{18}$$

where $P(\text{ Skill } | w_i)$ is the conditional probability that the competition is a game of skill given the observed win fraction of the ith player, w_i ; $P(w_i | \text{ Skill })$ is the probability of observing w_i if the competition is a game of skill (the *likelihood*); and $P(\text{Skill})$ is the *prior probability* that the competition is a game of skill. The denominator $P(w_i)$ is the probability of observing w_i regardless of contest type (the evidence) and is given by

$$P(w_i) = P(w_i | \text{ Luck }) P(\text{ Luck }) + P(w_i | \text{ Skill }) P(\text{ Skill }).$$

If we take z_i to be the number of wins and N_i to be the number of entries of the ith player (note that N_i, the number of entries, may be greater than n_i, the number of games, since players may submit more than one entry per game), $P(w_i | \text{ Luck })$ is equivalent to the probability distribution one would observe in a coin flipping competition and is given by the binomial distribution,

$$P(w_i|\text{ Luck }) = P_{\text{bin}}(z_i = N_i w_i) = \begin{pmatrix} N_i \\ z_i \end{pmatrix} \phi^{z_i}(1-\phi)^{N_i - z_i}$$

where $\phi = 1/2$ and $\begin{pmatrix} N_i \\ z_i \end{pmatrix}$ is the binomial coefficient.

For a game in which the outcome is determined solely by skill, players can, in theory, be ranked according to their skill level. In any match-up, the higher ranked player wins the contest. Hence the best player wins all of their games, the next best wins all but one, etc., and the probability distribution $P(w_i|\text{ Skill })$ is uniform:

$$P(w_i|\text{ Skill }) = P_{\text{uni}}(z_i = N_i w_i) = 1/N_i.$$

Finally, since there are only two possible outcomes in this formulation (i.e. the sample space is $\{\text{Luck}, \text{Skill}\}$), $P(\text{Skill}) = 1 - P(\text{Luck})$.

Starting with a uniform prior of $P(\text{Luck}) = P(\text{Skill}) = 1/2$, we iteratively evaluate $P(w_i|\text{ Skill })$ using Equation (18) and the observed data $\{z_i, N_i\}$ for a randomly picked player until the probability, $P(\text{Skill})$ converges (see Figure 5, Left) to one or zero. As before, we consider different populations defined by the minimum number of games per player. For all four sports, we draw 200 random samples from the relevant population and iteratively compute $P(w_i|\text{Skill})$. We record the outcome of this exercise (a one or zero), and repeat the process fifty times. We then average over all fifty outcomes to compute a mean $\overline{P}(\text{Skill})$. This process is repeated 50 times which allows us to compute not only the mean of $P(\text{Skill})$ but also the standard deviation of our computed mean as represented by the error bars in Figure 5 (Right). In all cases we find that $\overline{P}(\text{Skill})$ converges to one if the minimum number of games per player is sufficiently large. Hence for each sport there exists a transition game number, N_{TG} above which the outcome of the game is definitively determined by skill, i.e. for playing populations in which every player has played more games than N_{TG}, $\overline{P}(\text{Skill})$ converges to one. For all four sports in our FanDuel dataset, $\overline{P}(\text{Skill})$ is always greater than 1/2; furthermore for all NBA populations, $N_{TG} \approx 1$ suggesting that skill is always the dominant factor in determining the outcome of these contests regardless of how many games are played.

Although one should not necessarily expect a one-to-one mapping between the Bayesian approach and the previous analysis, we can check that trends and features are consistent across both methods. First, the ordering is as we would expect with fantasy basketball being the most skill-based and fantasy-hockey being the closest of the four to chance. More quantitatively, we can estimate the minimum number of games required to cross over into the skill-dominated region from Figure 4, and compare those numbers with N_{TG} from Figure 5 (Right). If we take $R^* = 1/2$ as the critical cross-over value, we would expect to see $N_{TG,MLB} \approx 20$, $N_{TG,NBA} \approx 5$, $N_{TG,NFL} \approx 25$, and $N_{TG,hockey} \approx 40$. These numbers for NBA, MLB, and NHL are surprisingly close across the two methods; for NFL, the R^* calculation is slightly more conservative, overestimating N_{TG} by approximately 10 games.

4. Perspectives on the relative role of skill and chance in games and other activities. The outcomes of these tests leave no doubt that skill plays a role in the outcome in fantasy sports competitions. However, it is useful to add perspective to these results by considering the relative role of skill and luck in the context of other

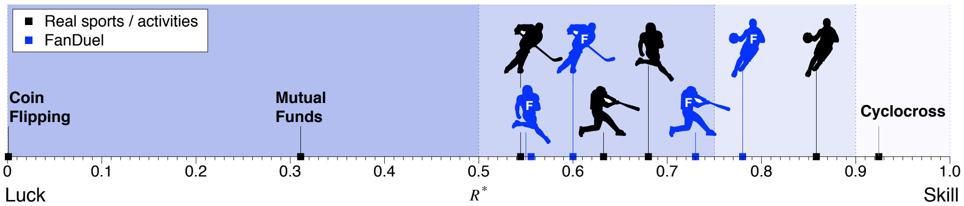

FIG. 6. *R^* values computed for fantasy sports, real sports, cyclocross racing, coin flipping, and mutual funds. Games of pure luck lie on the left and games of pure skill lie on the right. Blue sports icons represent fantasy sports and black sports icons represent professional sports.*

activities. The third test, persistence, can easily be applied to data from a variety of activities as shown in Figure 6 which quantifies R^* values for the four fantasy sports discussed in this study; real MLB, NBA, NFL, and NHL athletic competitions; cyclocross racing; coin flipping; and mutual funds. The R^* values for real MLB, NBA, NFL, and NHL athletic competitions were computed using publicly available data from the past five seasons (2010-2014). Each point in the $\hat{p}_i$ versus $\hat{q}_i$ plot represents the win fraction for the first half versus the win fraction for the second half of a single season for a particular professional sports team (hence each team is represented by five points on the p-q scatter plot corresponding to the five seasons we considered).

As one might expect, we find basketball at the skill end of the spectrum since there are many games in a basketball season and many scoring opportunities per game. Hence small differences in skill are amplified over the course of a season and strong teams tend to come out on top. At the other end of the sports cluster, we find hockey which typically has a small number of goals per game hence one "lucky shot" can make a big difference. While the R^* values for most of the sports and fantasy sports are consistent with our expectations, there is one point that is somewhat puzzling. Note that each sport and fantasy sport pair are relatively close to one another. The exception is professional football. According to the computed R^* value, professional football contests are largely determined by skill. This is somewhat surprising since there are only 16 games per team in the NFL season and the number of scoring opportunities per game is limited (compared to e.g. basketball). We currently have no explanation for this and we leave it as a puzzle for future investigation.

To compute the R^* values associated with cyclocross racing, we considered the finishing places of the top 30 performers using publicly available data from crossresults.com for 2015. For each athlete, $\hat{p}_i$ and $\hat{q}_i$ values were computed using the average finishing place for the first and second half of the season i.e. for each race, first place corresponds to 1, second place corresponds to 2, etc. Note that each athlete participates in a different set of races (although elite performers tend to all participate in a subset of key events). These $\hat{p}_i$ and $\hat{q}_i$ values were again rotated and used to compute R^* – taking care to include $E[S]$ which is not zero in this case – as shown in Figure 6. The coin flipping data point was computed using a simulation of a population of 100 players flipping 100 coins. The average R^* value over 100 trials – which not surprisingly is approximately zero – is shown in Figure 6.

As we have seen in the fantasy sports data, the measured value of R^* depends on the number of games associated with each player, n_i. For real sports it is easy to select a representative value of n_i since each team plays the same number of games in a season. For fantasy sports, the choice is less obvious since each player plays in a different number of contests. To estimate the characteristic number of games per player, we seek an appropriate average. Because the number of players who have played n_i games decays approximately exponentially (as shown in Figure 4), computing a simple average not ideal since the signal is swamped by players who have only played a few games. To address this we computed a logarithmically weighted average:

$$\overline{n} = \sum_{j=1}^{G_{max}} \frac{j \log(m_j)}{\log(m_j)}.$$

Here G_{max} is the maximum number of games played by any player in the playing population and m_j is the number of players who have played j games. Using this weighed average to compute a characteristic number of games played in each sport, we find $\overline{n}_{MLB} = 110$, $\overline{n}_{NBA} = 82$, $\overline{n}_{NFL} = 34$, and $\overline{n}_{NHL} = 60$. These values were used to compute the values of R^* for the fantasy sports shown in Figure 6.

Finally, we consider mutual funds. Mutual funds – investment programs controlled by (perhaps skillful) managers – have long been considered strategies that can beat the market while limiting risk. Savvy investors are constantly evaluating whether skilled money managers produce sufficient returns to justify their cost. Previous studies have found that chance certainly plays a role in manager performance, but placement on the skill-luck spectrum ranges from predominantly chance [8] to a more balanced skill-chance split [18]. In the context of the current study, a mutual fund manager – similar to a player in fantasy sports – must decide how to allocate funds to achieve optimal performance by identifying value in an open market. A fantasy sports player picks athletes who are projected to yield the most points relative to their price; a mutual fund manager invests in companies that they project to yield returns higher than their trading price. Hence we can compute R^* for mutual fund managers using market-adjusted mutual fund performance data (i.e. the performance of each fund was evaluated relative to the performance of the overall market) from Wharton Research Data Services (WRDS) from the past ten years (2005-2015). Here $\hat{p}_i$ and $\hat{q}_i$ for each mutual fund were computed for first and second half of the year, respectively, for each of the 10 years. We considered 44,938 distinct mutual funds with a total of 307,471 "entries." The results are shown in Figure 6. This estimate yields a diplomatic answer that splits the difference between previous chance-dominant calculations ($R^* \approx 0$) [8] and previous skill-chance balanced results ($R^* \approx 0.5$) [18].

4.1. Speculations on game design. To some extent – as beautifully articulated by Clauset *et al.* – it is the artful balance of skill and chance that makes sports so compelling:

> "On one hand, events within each game are unpredictable, suggesting that chance plays an important role. On the other hand, the athletes are highly skilled and trained, suggesting that differences in ability are fundamental. This tension between luck and skill is part of what makes these games exciting for spectators ..." [4]

Striking this balance is essential in the design of any competition. In any type of game there are a number of strategies a game designer can adopt to adjust the relative importance of luck in the outcome. First let us consider the effect of the distribution of skill within the playing population. To illustrate the importance of skill distribution consider a professional golfer playing against a novice versus two professionals (or two novices) playing each other. In the first case the outcome is a near certainty since the skill of the professional will dominate. In the second case, if the ability of the two players is similar, skill is no longer a distinguishing characteristic and the relative importance of chance in the outcome increases [20, 21]. Hence tournaments that are divided up into classes of different skill levels (e.g. having beginners play in a separate pool) are likely to have a larger element of luck than those in which everyone plays in the same pool.

The second game parameter that games designers may choose to adjust is the number of contests per player. Calculating the overall win probability in a best of seven series given the win probability of an individual game is a common exercise assigned in elementary probability courses and it is well-known that the role of skill is amplified through multiple contests. In the words of Levitt *et al.* "Even tiny differences in skill manifest themselves in near certain victory if the time horizon is long enough" [16]. Hence perhaps the simplest way to increase the role of skill in a contest is to increase the number of games per player in the competition.

Finally, game designers can address the balance of skill and chance head-on by addressing the role of chance as reflected in the rules of the game. For example, in a fantasy sports salary cap game, one of the parameters that can be tuned by game designers is player pricing. It is interesting to note that more accurate pricing algorithms push games towards the luck end of the spectrum. Consider the extreme case of perfect pricing in which the price of the player exactly mirrors their expected payoff. In this case, there is no strategy in assembling a line-up (other than to get as close to the salary cap as possible) and the outcome of the fantasy game is determined purely by luck. However, as the pricing becomes less accurate (i.e. less reflective of the expected payoff), skilled fantasy players can capitalized on undervalued players. Hence, to increase the role of skill in a fantasy competition, game designers could either add random noise to their pricing algorithms or increase points awarded for less-frequent, larger-variance events such that "perfect pricing" is inherently more difficult.

Acknowledgments. The authors would like to thank FanDuel for providing access to their data and for partial funding of the research. FanDuel did not exercise any editorial control over this paper. The authors would also like to thank two anonymous referees for their creative and insightful suggestions with led to the section on Bayesian inference and the "quitting boundary layer" analysis.

Appendix A. FanDuel Data, Rules, and Scoring. The number of players and the range of entries per player in each group in the expected payoff calculation is summarized in Table 4. Roster positions and scoring for the FanDuel 2014 season are summarized in Tables 1, 6 and 5 respectively.

		Grp 1	**Grp 2**	**Grp 3**	**Grp 4**	**Grp 5**
MLB 2013/14	average w_i	0.4741	0.5138	0.5706	0.5991	0.602
	# plrs	28,027	2,803	280	28	5
	range of entries	6–141	142–1,552	1,553–6,990	6,991–14,088	14,089–24,390
MLB 2014/15	average w_i	0.4626	0.4967	0.5598	0.5853	0.5851
	# plrs	66,556	6,656	666	66	9
	range of entries	6–132	132–1,566	1,566–10,512	10,513–28,632	28,631–60,205
NBA 2013/14	average w_i	0.4306	0.5096	0.5678	0.5891	0.6312
	# plrs	21,635	2,163	216	22	4
	range of entries	6–160	160–1,485	1,485–6,301	6,302–14,279	14,280–64,292
NBA 2014/15	average w_i	0.4319	0.4891	0.5764	0.612	0.6141
	# plrs	88,435	8,844	884	89	11
	range of entries	6–124	124–1,476	1,476–9,666	9,667–26,636	26,636-41,621
NFL 2013/14	average w_i	0.5253	0.54	0.5613	0.5715	0.6336
	# plrs	32,633	3,264	326	33	5
	range of entries	6–42	42–248	249–1,853	1,854–6,726	6,727–49,662
NFL 2014/15	average w_i	0.5102	0.5429	0.596	0.6414	0.6542
	# plrs	117,582	11,758	1,176	117	15
	range of entries	6–43	43–311	312–2,868	2,869–13,239	13,240–26,478
NHL 2013/14	average w_i	0.5155	0.526	0.549	0.5642	0.5914
	# plrs	4,625	463	46	5	2
	range of entries	6–103	103–610	611–1,871	1,872–4,377	4,378–16,787
NHL 2014/15	average w_i	0.5062	0.5319	0.5652	0.5757	0.5757
	# plrs	16,577	1,657	166	17	3
	range of entries	6–108	108–918	919–5,964	5,965–12,110	12,111–21,214

TABLE 4
Summary of average win fraction, number of players, and the range of number of entries per player in each group as computed in section 2.1.

MLB		**NBA**		**NFL**		**NHL**	
W	4			Sacks	1	Wins	3
ER	−1			Fumbles recovered	2	Goals against	−1
SO	1			Return touchdowns	6	Saves	0.2
IP	1			Extra point return	2	Shutouts	2
				Safeties	2		
				Blocked punt/kick	2		
				Interceptions made	2		
				0 points allowed	10		
				1-6 points allowed	7		
				7-13 points allowed	4		
				14-20 points allowed	1		
				28-34 points allowed	−1		
				35+ points allowed	−4		

TABLE 5
FanDuel 2013 and 2014 scoring (defense).

MLB		**NBA**		**NFL**		**NHL**	
1B	1	3-pt FG	3	Rushing yards made	0.1	Goals	3
2B	2	2-pt FG	2	Rushing touchdowns	6	Assists	2
3B	3	FT	1	Passing yards	0.04	Plus/minus	1
HR	4	Rebound	1.2	Passing touchdowns	4	Penalty minutes	0.25
RBI	1	Assist	1.5	Interceptions	−1	Powerplay points	0.5
R	1	Block	2	Receiving yards	0.1	Shots on goal	0.4
BB	1	Steal	2	Receiving touchdowns	6		
SB	2	Turnover	−1	Receptions	0.5		
HBP	1			Kickoff return TDs	6		
Out	−0.25			Punt return TDs	6		
				Fumbles lost	−2		
				2-point conversions scored	2		
				2-point conversion passes	2		
				FG from 0-39 yds	3		
				FG from 40-49 yds	4		
				FG from 50+ yds	5		
				Extra point conversions	1		

Table 6
FanDuel 2013 and 2014 scoring (offense).

REFERENCES

[1] To request access to these data sets for academic purposes, please contact fdlegal@fanduel.com.
[2] *Fantasy sports trade association: Industry demographics*. http://fsta.org/research/industry-demographics/. Accessed: 10-15-2015.
[3] *Unlawful internet gambling enforcement act*, 2006. title 31, section 5362, 1.E.ix.
[4] A. Clauset, M. Kogan, and S. Redner, *Safe leads and lead changes in competitive team sports*, Phys Rev E, 91 (2015), p. 062815.
[5] R. Croson, P. Fishman, and D. Pope, *Poker superstars: Skill or luck?*, Chance, 21 (2008), p. 25, doi:doi:10.1007/s00144-008-0036-0.
[6] M. Dreef, P. Borm, and B. van der Genugten, *Measuring skill in games: several approaches discussed*, Math Meth Oper Res, 59 (2004), pp. 375–391, doi:10.1007/s001860400347.
[7] L. Epstein and A. Martin, *An Introduction to Empirical Legal Research*, Oxford University Press, 2014.
[8] E. F. Fama and K. R. French, *Luck versus skill in the cross-section of mutual fund returns*, The Journal of Finance, 65 (2010), pp. 1915–1947.
[9] A. Gabel and S. Redner, *Random walk picture of basketball scoring*, J. Quant. Anal. Sports, 8 (2012).
[10] A. Heuer, C. Muller, and O. Rubner, *Soccer: Is scoring goals a predictable poissonian process?*, Europhys. Lett., 89 (2010), p. 38007.
[11] O. W. Holmes, *The path of the law*, Harv. L. Rev., 10 (1897), p. 457.
[12] W. Hurley, *Overtime or shootout: Deciding ties in hockey*, in ASA-SIAM Series on Statistics and Applied Mathematics: Anthology of Statistics in Sports, J. Albert, J. Bennett, and J. J. Cochran, eds., SIAM, 2005, ch. 25, pp. 193–196.
[13] D. Jackson and K. Mosurski, *Heavy defeats in tennis: Psychological momentum or random effect?*, in ASA-SIAM Series on Statistics and Applied Mathematics: Anthology of Statistics in Sports, J. Albert, J. Bennett, and J. J. Cochran, eds., SIAM, 2005, ch. 42, pp. 303–310.
[14] D. P. Kiley, A. J. Reagan, L. Mitchell, C. M. Danforth, and P. S. Dodds, *Game story space of professional sports: Australian rules football*, Phys. Rev. E, 93 (2016), p. 052314.
[15] S. D. Levitt and T. J. Miles, *The role of skill versus luck in poker evidence from the world series of poker*, Journal of Sports Economics, 15 (2014), pp. 31–44, doi:10.1177/1527002512449471.
[16] S. D. Levitt, T. J. Miles, and A. Rosenfield, *Is texas hold 'em a game of chance? a legal and economic analysis*, 101 Geo. L.J., 581 (2012–2013), pp. 581–636.
[17] G. R. Lindsey, *The progress of the score during a baseball game*, in ASA-SIAM Series on Statistics and Applied Mathematics: Anthology of Statistics in Sports, J. Albert, J. Bennett, and J. J. Cochran, eds., SIAM, 2005, ch. 16, pp. 119–144.
[18] A. Mauboussin and S. Arbesman, *Differentiating skill and luck in financial markets with streaks*, 2011, doi:http://dx.doi.org/10.2139/ssrn.1664031. Available at SSRN: https://ssrn.com/abstract=1664031.
[19] D. G. Morrison and M. U. Kalwani, *The best nfl field goal kickers: Are they lucky or good?*, in ASA-SIAM Series on Statistics and Applied Mathematics: Anthology of Statistics in Sports, J. Albert, J. Bennett, and J. J. Cochran, eds., SIAM, 2005, ch. 7, pp. 45–52.
[20] F. Mosteller and C. Youtz, *Professional golf scores are poisson on the final tournament days*, in 1992 Proceedings of the Section on Statistics in Sports, American Statistical Association, 1992, pp. 39–51.
[21] F. Mosteller and C. Youtz, *Where eagles fly*, Chance, 6 (1993), pp. 37–42.
[22] P. K. Newton and K. Aslam, *Monte carlo tennis*, SIAM Review, 48 (2006), pp. 722–742.
[23] H. V. Ribeiro, S. Mukherjee, and X. H. T. Zeng, *Anomalous diffusion and long-range correlations in the score evolution of the game of cricket*, Phys. Rev. E, 86 (2012), p. 022102.
[24] H. S. Stern, *A brownian motion model for the progress of sports scores*, in ASA-SIAM Series on Statistics and Applied Mathematics: Anthology of Statistics in Sports, J. Albert, J. Bennett, and J. J. Cochran, eds., SIAM, 2005, ch. 34, pp. 257–263.
[25] A. Tversky and T. Gilovich, *The cold facts about the 'hot hand' in basketball*, Chance, 2 (1989), pp. 16–21.
[26] R. P. van Loon RJD, M. van den Assem, and D. van Dolder, *Beyond chance? the persistence of performance in online poker*, PLoS ONE, 10 (2015), p. e0115479, doi:doi:10.1371/journal.pone.0115479.
[27] R. L. Wardrop, *Simpson's paradox and the hot hand in basketball*, in ASA-SIAM Series on Statistics and Applied Mathematics: Anthology of Statistics in Sports, J. Albert, J. Bennett, and J. J. Cochran, eds., SIAM, 2005, ch. 22, pp. 175–179.

[28] G. Yaari and G. David, *The hot (invisible?) hand: Can time sequence patterns of success/failure in sports be modeled as repeated independent trials*, PLoS ONE, 6 (2011), p. e24532.