Journal of Sports Sciences

ISSN: 0264-0414 (Print) 1466-447X (Online) Journal homepage: http://www.tandfonline.com/loi/rjsp20

Predicting coin flips: using resampling and hierarchical models to help untangle the NHL's shoot-out

Michael J. Lopez & Michael Schuckers

To cite this article: Michael J. Lopez & Michael Schuckers (2016): Predicting coin flips: using resampling and hierarchical models to help untangle the NHL's shoot-out, Journal of Sports Sciences, DOI: 10.1080/02640414.2016.1198046

To link to this article: http://dx.doi.org/10.1080/02640414.2016.1198046

Published online: 04 Jul 2016.

JOURNAL OF SPORTS SCIENCES, 2016
http://dx.doi.org/10.1080/02640414.2016.1198046

Predicting coin flips: using resampling and hierarchical models to help untangle the NHL's shoot-out

Michael J. Lopez[a] and Michael Schuckers[b]

[a]Department of Mathematics, Skidmore College, Saratoga Springs, NY, USA; [b]Department of Mathematics, Computer Science, and Statistics, St. Lawrence University, Canton, NY, USA

ABSTRACT

Roughly 14% of regular season National Hockey League games since the 2005–06 season have been decided by a shoot-out, and the resulting allocation of points has impacted play-off races each season. But despite interest from fans, players and league officials, there is little in the way of published research on team or individual shoot-out performance. This manuscript attempts to fill that void. We present both generalised linear mixed model and Bayesian hierarchical model frameworks to model shoot-out outcomes, with results suggesting that there are (i) small but statistically significant talent gaps between shooters, (ii) marginal differences in performance among netminders and (iii) few, if any, predictors of player success after accounting for individual talent. We also provide a resampling strategy to highlight a selection bias with respect to shooter assignment, in which coaches choose their most skilled offensive players early in shoot-out rounds and are less likely to select players with poor past performances. Finally, given that per-shot data for shoot-outs do not currently exist in a single location for public use, we provide both our data and source code for other researchers interested in studying shoot-out outcomes.

ARTICLE HISTORY
Accepted 26 May 2016

KEYWORDS
Hockey; shoot-outs; hierarchical models; Bayesian models; reproducibility

Introduction

Following the locked out 2004–05 regular season, the National Hockey League instituted a shoot-out to determine winners of regular season games that finished overtime still tied. Shoot-outs in hockey take a similar form to penalty kicks in association football (soccer) matches. In the NHL's adaptation, both teams take alternating penalty shots three times. If the teams are still tied after those three rounds, then the teams complete single rounds until one team scores and the other does not.

To ensure that the shoot-out was taken seriously, the NHL also updated its point system for the 2005–06 season. As a result, teams victorious in the shoot-out are awarded two points towards league standings, identical to the number awarded for a regulation or overtime win. Teams can also earn a single point for an overtime or shoot-out loss.

Roughly 1 in 7 NHL games have reached a shoot-out since its implementation, which, in an 82-game season, equates to an average of 11 or 12 points a year for each team that are decided by the post-overtime competition. Given the league's point system, shoot-out performance can make or break a team's season, as standings often yield little separation between the top play-off qualifiers and teams not participating in the play-offs. For example, San Jose (No. 12 finisher, 89 points) finished within 15 points of Nashville (No. 3, 104 points) in the final 2014–15 standings in the 14-team Western Conference. Perhaps unsurprisingly, Seppa (2009) identified the shoot-out as a deciding factor in which teams make the play-offs. With NHL teams earning a disproportionate amount of the profit that comes in play-off ticket sales (Leeds & Von Allmen, 2004), the shoot-out also has financial ramifications.

While the league's revenue has increased since the shoot-out's implementation (Stubits, 2014), critics of the shoot-out have called it a "gimmick" (Feschuk, 2014). Notably, Calgary Flames' president Brian Burke called it a "circus stunt … horse (bleep)" (Brough, 2014). Much of the frustration with respect to the shoot-out lies in its randomness, as most consider the outcome to be no different than a "coin flip" (Feschuk, 2014; Schuckers, 2009). But is it? Despite the shoot-out's consistent and powerful influence on the NHL, there is little published work on the factors that drive individual and team performance. Moreover, there is no centralised shot-by-shot data set for use; even the league's own website, www.nhl.com, computes only aggregated player statistics for the shoot-out. This stands in stark contrast to professional soccer, where the shoot-out has been studied extensively. See Pollard and Reep (1997), McGarry and Franks (2000), Jordet, Hartman, Visscher, and Lemmink (2007), Jordet (2009), Apesteguia and Palacios-Huerta (2010), Wood, Jordet, and Wilson (2015), to name a few, for examples of analyses for soccer.

Most analyses of NHL shoot-outs have come in the form of blog posts. Among these are three articles by Gabe Desjardins (writing as Hawerchuk). In the first of these, Desjardins looked at the distribution (in shots) of the shoot-out through March of 2009 and compared that to what would be expected given a Bernoulli model with success probability 33% for each shot (Desjardins, 2009a). Later, Desjardins wrote about comparing

CONTACT Michael Lopez mlopez1@skidmore.edu

past performance in the shoot-out to future performance for shooters (Desjardins, 2009b) and for goalies (Desjardins, 2009c). His conclusion based upon these position-specific analyses is that there is not much difference in the performance of the top players from the bottom players. Schuckers (2009) provided a simple Bayesian analysis of the first 5 years of the shoot-out in the NHL using a probit model that included a term for each shooter and each goalie, assuming a flat prior on each parameter. We build upon that analysis below. Schuckers' conclusion was that the shoot-out is randomness, a crapshoot. Eric Tulsky (aka Eric T.) (2009) used additional data and reached similar conclusions. More recently, Schuckers and Nelson (2013) looked at a team's choice to shoot first or second, finding no significant difference.

There is scant peer-reviewed literature on the subject. McEwan, Ginis, and Bray (2012) found statistically significant effects of pressure on performance. Specifically, the authors identified that visiting team players performed better in win-imminent situations (shooting to win) relative to home team players (an absolute difference of 7%), but that home team players performed significantly better in loss-imminent situations (trying to avoid a loss, an absolute difference of 9%). Jones (2013) looked at the impact of home advantage across a variety of sports, finding that in the first 4 years of the NHL shoot-out, there was no home advantage. In neither of these papers did the authors control for the individual talent levels among goaltenders or shooters. While mostly unrelated to our goals, Hurley (2005) took a statistical look at alternatives to the shoot-out, and Hurley (2008) analysed approximately a season and a half of the NHL shoot-out to model its length.

In contrast to hockey, there is a sizeable peer-reviewed literature on shoot-outs in soccer, which take the form of alternating penalty kicks. Pollard and Reep (1997) discuss shoot-outs as part of an in-depth statistical study of events in soccer, and the optimal order of players is analysed by McGarry and Franks (2000). Jordet et al. (2007) used modelling to suggest that there is evidence that shooters choke in high-pressure situations, the results of which were expanded upon by Jordet (2009). For example, goal rates on penalty kicks were 30% lower among kick takers that needed to score to tie the game (Jordet, 2009). Apesteguia and Palacios-Huerta (2010) found that teams shooting first in the shoot-out won more often (60% of the time), and that scoring rates increase in later rounds for teams that are leading and scoring rates decrease for teams that are behind. In fact, even team's uniform colour and "pre-penalty gaze" have been discussed as predictors of performance (Greenlees, Leyland, Thelwell, & Filby, 2008). We study some of these identical covariates to consider their impact on the scoring rates in hockey shoot-outs.

This manuscript makes four major contributions. First, we look at overall team and individual shoot-out performances to identify how players perform relative to chance alone. Second, we use advanced regression modelling strategies to identify if players perform better or worse under pressure, as well as to quantify and compare within shooter and within goalie variability. Our results suggest that, while it is difficult to distinguish between the performances of most players, there are certain shooters and a few goalies who have outperformed expectations by statistically significant margins. Meanwhile, there is no evidence that performance links to the shooter's pressure, based on if a goal is needed to extend the shoot-out, if a goal will end the shoot-out, or if the shoot-out will continue regardless of the shot outcome.

Third, we implement a set of simulations that identify how coaches select players for the shoot-out, as well as to identify the financial worth of the league's top players to their teams. We estimate that the league's best shooters and goalies have been worth more than half of a million American (US) dollars per year to their teams on shoot-out performance alone. Finally, given that per-shot data are missing in easily analysable format, as well as the increased awareness regarding the importance of reproducible scientific research (Peng, 2011), we provide all of our data and code so that future researchers can confirm our results while performing their own analyses.

Methods

Data collection

Using play-by-play game output provided by www.nhl.com, we collected information on the 1583 shoot-outs in the 10 full NHL seasons between 2005–06 and 2014–15. This contained 10,839 shot attempts during these seasons, 3604 (33.3%) of which were successfully converted by the shooter. There are 164 different goalies and 828 different shooters in these data. The shooter taking the shot (identified by name), his position (defence or forward), the goalie facing the shot (identified by name), whether or not the shot resulted in a goal (yes or no), the current score of the shoot-out (relative to the shooter, −2, −1, 0 or 1), the location of each team (home or away) and shot number (e.g. first, second and fifth) were all recorded.

We notate these variables as follows. Our outcome is $Goal_{ijk}$, whether or not shooter j scored against goalie k on attempt i, for $i = 1 \dots 10839$, $j = 1, \dots 828$ and $k = 1, \dots 164$. Let $Defence_i$ be an indicator for whether or not the shooter of shot i is a defender, and let $Visiting_i$ be an indicator for whether or not shot i is attempted by the visiting team. Given the literature on how pressure can impact soccer penalty kicks, we define each shot as follows.

$$Status_i = \begin{cases} WinImminent, & \text{if the shooter ends the game with a goal on shot } i \\ LossImminent, & \text{if the shooter needs a goal on shot } i \text{ to keep game alive} \\ Early, & \text{otherwise.} \end{cases} \tag{1}$$

Team analysis

We first contrast win percentage for home versus away teams. Home teams have been given the opportunity to shoot or defend first in every season since 2005–06. In the 2005–06 season, home teams were required to shoot second. If there has been a home advantage, assuming all other factors held constant, we expect the home team to win more than 50% of shoot-outs. A one-proportion z-test will be used to judge significance. Given evidence that shooting first or second matters in soccer (Apesteguia & Palacios-Huerta, 2010), we perform a similar analysis to judge if shoot-out order matters in hockey as far as team success. Finally, we will check how well a team's performance translates from one season to the next using a scatter plot, with the *x*-axis being the team's shoot-out win percentage in one year and the *y*-axis that team's win percentage in the following year.

Individual analysis

Judging shooter (or goalie) success by goal (save) percentage alone is difficult, given that many players have only participated in a few shoot-outs. For example, while there are 39 shooters who have scored on all of their attempts, none of these skaters have attempted more than two shots. As a result, it is difficult to contrast their perfect records to players having a substantially larger number of attempts.

In place of ranking players on goal or save percentages alone, funnel plots allow visualisation of shooter and goalie performance while factoring in their relative sample sizes (Spiegelhalter, 2005). We briefly describe construction of a funnel plot for goalies; our plot for shooters is similarly constructed.

Our goalie funnel plot contrasts each goalie's career save percentage against his number of attempts. At the centre of the *y*-axis lies the population save rate, $\lambda = 0.668$, which, under a hypothesis that all goalies are identical, represents the expected value of each goalie's individual rate. Two confidence limits are shown, representing the bounds for individual performances lying two (95% limits) or three (99.8%) standard deviations above or below λ. As in Spiegelhalter (2005), we use the exact limits from the inverse Binomial distribution, which are preferred over a normal approximation with smaller sample sizes.

If shoot-outs are a random outcome with save probability λ, we expect roughly 95% and 99.8% of goalies to fall within two and three standard deviations of λ. If there are more players than we expect beyond these boundaries, it would provide evidence that there is some inherent skill in saving shoot-out attempts. Thus, we check for the relative number of goalies falling beyond these boundaries in the funnel plots as one measure of the shoot-out's true randomness. For shooters, funnel plots are centred at $1 - \lambda$. Both funnel plots are presented in the results section of the manuscript.

Shooter selection bias

The results of McEwan et al. (2012) suggest that NHL shooter performance varies when a goal is needed either to extend or to end the shoot-out. However, a critical component of the shoot-out is that while goalies almost universally remain in net for the duration of the shoot-out, coaches can place a shooter in any round until each of a team's shooters has one attempt apiece. As a result, there is the potential that if the league's best shooters are always shooting first, the association between a shot's pressure and decreased shooter performance could be accounted for by changes in shooter talent.

Table 1 shows the shoot-out rounds of the league's five best shooters, as will be identified later. More than 92% of these shooters' attempts have been in rounds one or two, including 80 attempts from Jonathan Toews, who has never taken a shot after round two.

Given that shoot-outs can end as early as round 2 (two goals from one team and blanks from its opponent), it is understandable why coaches would deploy their best shooters early. However, such an allocation makes it difficult to look at player performance under pressure, given that the types of shooters that shoot early (with generally less pressure) are likely more talented than those that shoot later (generally more pressure). Perhaps a preferable question would be to contrast shooter performance under pressure among shooters who have taken a sufficient number of attempts in all rounds and under all pressures; this strategy is implemented next.

Table 1. Round deployment among top shooters.

Shooter	Round 1	Round 2	Round 3	Round 4+
Nielsen	48	23	0	1
Toews	71	9	0	0
Oshie	41	9	7	2
Kozlov	36	32	9	3
Christensen	47	2	2	3
Total	243	75	17	9

Models of NHL shoot-out outcomes

To simultaneously account for several variables that may be concurrently impacting shoot-out rates, as well as the dependency within individual shooter and goalie outcomes, we propose multiple modelling frameworks.

Hierarchical mixed models

To account for the selection bias in the allocation of the best shooters to earlier shoot-out rounds, we will use two samples from our data. The first, the full sample, will contain all shots, while the second, identified as the reduced sample, will include only the attempts by shooters with at least five attempts in each of the first three shoot-out rounds. Let $n_j = 828$ and $n_k = 164$ and $n_{j'} = 88$ and $n_{k'} = 156$ be the number of shooters and goalies in the full and reduced samples, respectively.

Let $\mathbf{X_i}$ be a design matrix including fixed effects for $Status_i$ (*WinImminent* and *LossImminent* attempts compared to no pressure ones), $Visiting_i$ and $Defence_i$. To account for the correlation in $Goal_{ijk}$ between attempts made by the same shooter or against the same goalie, let δ_j, $\delta_{j'}$, δ_k and $\delta_{k'}$ be shooter- and goalie-specific random effects. These correspond to each shooter and each goalie in the full and reduced samples, respectively.

We propose three generalised linear mixed (GLM) models, or subsets thereof, where, letting $\text{logit}(p) = \log(\frac{p}{1-p})$,

$$\text{logit}(P(Goal_{ijk} = 1)) = \alpha + \beta_T \mathbf{X_i} \quad (2)$$

$$\text{logit}(P(Goal_{ijk} = 1)) = \alpha + \beta_T \mathbf{X_i} + \delta_j + \delta_k \quad (3)$$

$$\text{logit}(P(Goal_{ijk} = 1)) = \alpha + \beta_T \mathbf{X_i} + \delta_{j'} + \delta_{k'} \quad (4)$$

where $\delta_j \sim N(0, \tau_j^2)$, $\delta_k \sim N(0, \tau_k^2)$, $\delta_{j'} \sim N(0, \tau_{j'}^2)$ and $\delta_{k'} \sim N(0, \tau_{k'}^2)$.

Model (2) features a set of fixed effects but no random effects, while Model (3) includes random effects for both shooter and goalie. Model (4) uses only players in the reduced set of shots. If there is large variability between shooter performances after accounting for X, we would expect large τ estimates. We can also use the estimated τ terms to contrast the relative amount of variability between player performances.

Bayesian model of NHL shoot-out outcomes

Similar to the work of the *GLM* model above, we also fit a hierarchical Bayesian model to assess the shoot-out. Our model followed a similar structure to that proposed by Albert and Chib (1993), where

$$Goal_{ijk} \sim Bernoulli(z_{ijk}),$$

$$z_{ijk} \sim TruncMVN(\boldsymbol{\theta}^T X_i, \mathbf{1}),$$

$$\boldsymbol{\theta}^T = (\alpha, \boldsymbol{\delta}_g, \boldsymbol{\delta}_s, \beta)^T \sim MVN(0, \tau^T \mathbf{I}),$$

$$\tau_\alpha^2 \sim \text{Inverse } \chi^2(5),$$

$$\tau_g^2 \sim \text{Inverse } \chi^2(5),$$

$$\tau_s^2 \sim \text{Inverse } \chi^2(5),$$

$$\tau_\beta^2 \sim \text{Inverse } \chi^2(5),$$

such that τ is a concatenation of the precision parameters with $\boldsymbol{\tau} = (\tau_\alpha^2, \tau_j^2 \mathbf{1}_j, \tau_k^2 \mathbf{1}_k, \tau_\beta^2 \mathbf{1}_4)^T$ where 1_t is a vector of ones of length t. Then, τ_α is the precision for the intercept term; τ_g is the precision for the goalie terms; τ_s is the precision for the shooter terms and τ_β is the precision for the other covariates in our model. We are thus assuming that there is a latent variable for each shot, z_{ijk}, which depends on j, k, $Status_i$, $Visiting_i$ and $Defence_i$. The TruncMVN represents a truncated multivariate normal distribution which is truncated to only positive values if $Goal_{ijk} = 1$ and only negative values if $Goal_{ijk} = 0$. The expected value of the latent variable, z_{ijk}, is determined by the product of the model coefficients, θ, and the design matrix $\mathbf{X}$, where $\mathbf{X}$ is the matrix of covariates with each row corresponding to X_i, the observed covariates for the ith shoot-out shot attempt. The precision of the z_{ijk} is unity, again following Albert and Chib (1993).

The model coefficients, denoted by θ, are a concatenation of the parameters for goalies, shooters and the other covariates in the model, namely *Status, Visiting* and *Defence*. The prior for these coefficients, $\boldsymbol{\theta}$, is a multivariate normal distribution with mean zero and precision matrix a vector τ times the identity matrix. We follow Gelman (2006) in using an inverse χ^2 for the prior distribution of our precision parameters, the τ's. Our choice of five degrees of freedom for the prior distribution of the τ is one that assumes some stability to the model precision but that can be "swamped" by the large number of observations here.

Consent for the data collection and analysis was given by the Institutional Review Board at Skidmore College. Statistical analysis was conducted using the *R* statistical software (R Core Team, 2015), and Models (2) through (4) were fit using the "lme4" package (Bates et al., 2015). All *R* code and data can be found at this URL: https://github.com/NHLshoot-outs/NHL-shoot-outs.

Results

Team level results

The road team won 52.2% of the NHL's 1583 shoot-outs, a percentage that is not significantly different than 0.5 ($P = 0.09$). Interestingly, road team success peaked in the 2010–11 season. Figure 1 shows the yearly win percentage for road teams in shoot-outs, as well as 95% confidence intervals for each yearly sample proportion. In all but the 2010–11 season, during which road teams won 91 of the 149 shoot-outs (61.0%), the road team win percentage has been within ± 5 points of 50%. Other than chance, we are unaware of any plausible explanation for the results from the 2010–11 season.

Altogether, 49.6% of teams shooting first eventually won the shoot-out, a percentage that did not noticeably vary by home or away team. Despite the decision to go first making no apparent difference on win likelihood, an overwhelming 77.6% of teams chose to go first.

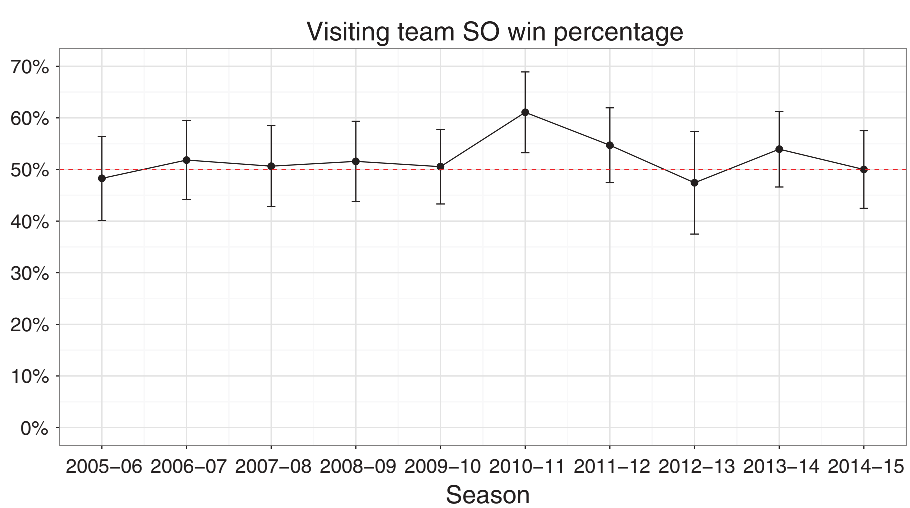

Figure 1. Road team yearly win percentage in the shoot-out, with 95% confidence limits.

Figure 2 shows a scatter plot of team shoot-out win percentage in one year versus its win percentage in the year prior, along with a locally weighted smoothed line and its 95% confidence limits. There is a slight positive increase to the line, indicating that teams winning a larger fraction of shoot-outs in one year may be more likely to win more shoot-outs in the next year. However, at all win percentages, the deviations from 50% are not significant, indicating that, by and large, team-level shoot-out ability does not meaningfully correlate from one season to the next.

Individual level results

Descriptive statistics and visualisations

Shooters in low-pressure situations successfully converted 33.9% of attempts, versus 32.5% of win imminent and 30.5% of loss imminent attempts. As judged by using a χ^2-test, these differences were significant (P = 0.03). Defencemen were only successful on 27.9% of attempts, compared to 33.6% for forwards (P = 0.003). However, coaches used defensemen on only 662 (6.1%) opportunities.

Figures 3 and 4 show funnel plots for skaters (all) and goalies (at least 10 attempts), respectively, along with confidence limits

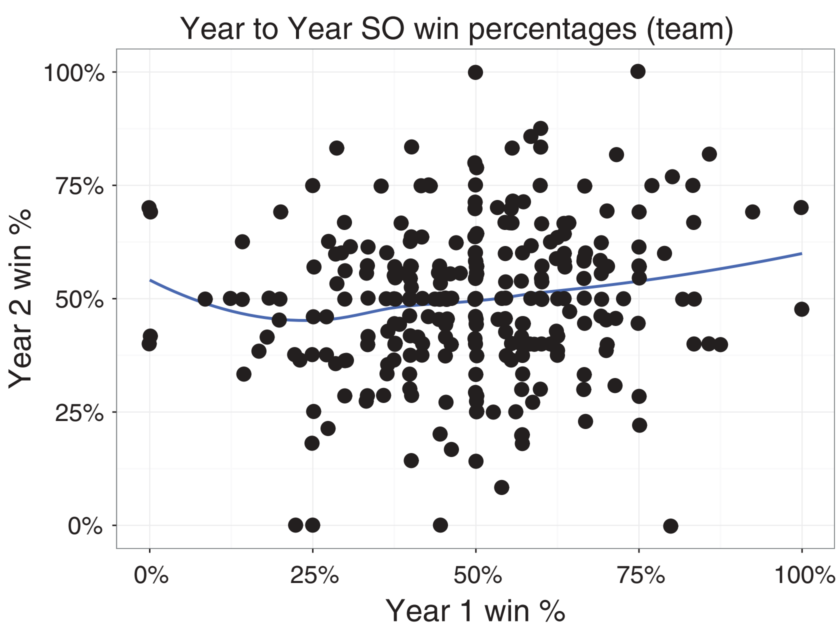

Figure 2. Year-to-year team winning percentage in the shoot-out, along with a locally weighted smoother and its 95% confidence limits.

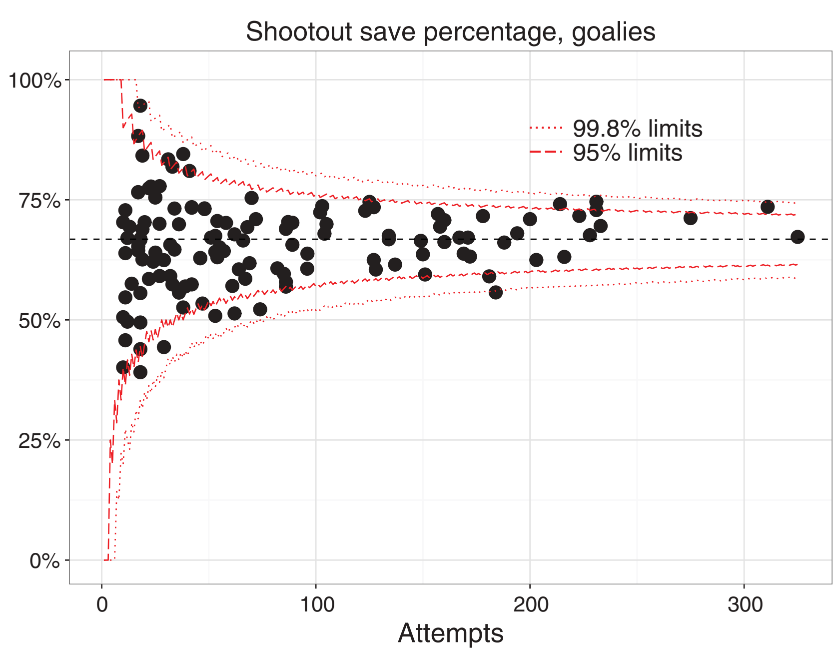

Figure 3. Goalie attempts and save rates, with dotted lines representing 95% and 99.8% confidence limits for individual goalie performance.

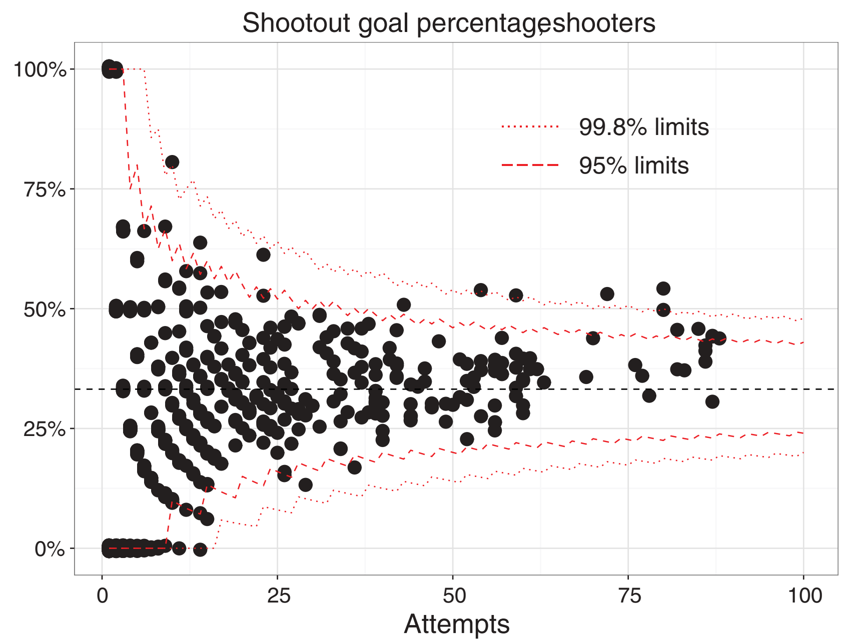

Figure 4. Shooter attempts and goal rates, with dotted lines representing 95% and 99.8% confidence limits for individual shooter performance.

for individual performances that lie two or three standard deviations above or below the league-wide averages.

In both of Figures 3 and 4, the majority of players are clustered within the confidence limits. However, particularly with shooters in Figure 4, there are a handful of players who are outperforming expectations, lying more than two or three standard deviations above the league average. Among the 15 shooters with at least 75 shoot-out attempts, 6 (40%) lie above the 97.5th percentile and 2 (13%) lie above the 99.8th percentile, at least with respect to our expectations under league-wide randomness. Of the 11 goalies with at least 200 attempts, 3 (27%) lie above the 97.5th percentile. This suggests, particularly among shooters, that there is variability in performance that exceeds what we would expect due to chance alone.

Relative to league-wide expectations, as well as their sample sizes, the league's best shooters are Eric Christensen (54 attempts, 53.7% success rate), Franz Nielsen (72, 52.8%), TJ Oshie (59, 52.5%), Victor Kozlov (80, 53.8%) and Jonathan Toews (80, 50%). Using a similar strategy, Henrik Lundvquist (311, 73.5% save percentage) rates as the league's most impressive netminder in shoot-outs, while Nicklas Backstrom (185, 54.1%) rates as its worst.

Regression model results

Following the advice of Gelman, Pasarica, and Dodhia (2002), and in place of tables with coefficient estimates, we use whisker plots of the estimated coefficients from Models (2)–(4), along with their 95% confidence intervals, as shown in Figure 5. Coefficients are presented on the log-odds scale, such that the positive estimates are associated with increased goal likelihood. (These estimates are available upon request, and also appear in the *R* code output.)

In Model (2), there is strong evidence that shooters perform significantly worse in loss imminent situations (Odds Ratio (OR) 0.85, 95% CI 0.74–0.96). This is shown in Figure 5 by an error bar for "Loss Imminent" situations that does not include 0. However, after accounting for the individual talent of goalies and shooters, as in Models (3) and (4), there is limited to no evidence that shooters are any better or worse under pressure. In the model fit on the reduced subset of the data, estimates for shots taken under both pressures are near 0.

Using the output from Model (2), the odds of a goal on shots by a member of the visiting team are about 8% higher (OR 1.08, 95% CI 1.00–1.18) than for home team attempts. Coefficient estimates for *Visiting* are robust to model specification.

The estimated random effect terms from Models (3) and (4) are $\hat{\tau}_j = 0.191$, $\hat{\tau}_k = 0.178$, $\hat{\tau}_{j'} = 0.104$, and $\hat{\tau}_{k'} = 0.073$. This suggests that the variability in shoot-out skill between shooters (j) is slightly larger than that between goalies (k). Further, the overall magnitudes of these estimates are far from zero, suggesting non-random variability applicable to both positions. Using the estimated shooter and goalie intercepts from Model (3), Victor Kozlov, Jonathan Toews, Frans Nielsen and TJ Oshie rank as the league's best shooters, with Michael Ryder, Tomas Plekanec, Clarke MacArthur and Martin Havlat ranking as its worst. Marc-Andre Fluery and Henrik Lundqvist rank as the league's best goalies (lowest random effects), with Nicklas Backstrom and Martin Biron ranking as the worst.

By and large, the results of the Bayesian approach are similar to those found by using frequentist methods. Using the model given in the previous section, we ran a single Markov chain Monte Carlo with a burn-in of 1000 iterations followed by 10,000 iterations where we thinned the chain by taking every 10th iteration. Thus, we have 1000 draws from the joint posterior distribution of the model parameters. Each of the posterior 99% credible intervals for the covariates in the Bayesian model (5), *Defence*, *Visiting* and *Status*, includes 0. This suggests that the a posteriori probability of these variables being associated with shoot-out outcomes is

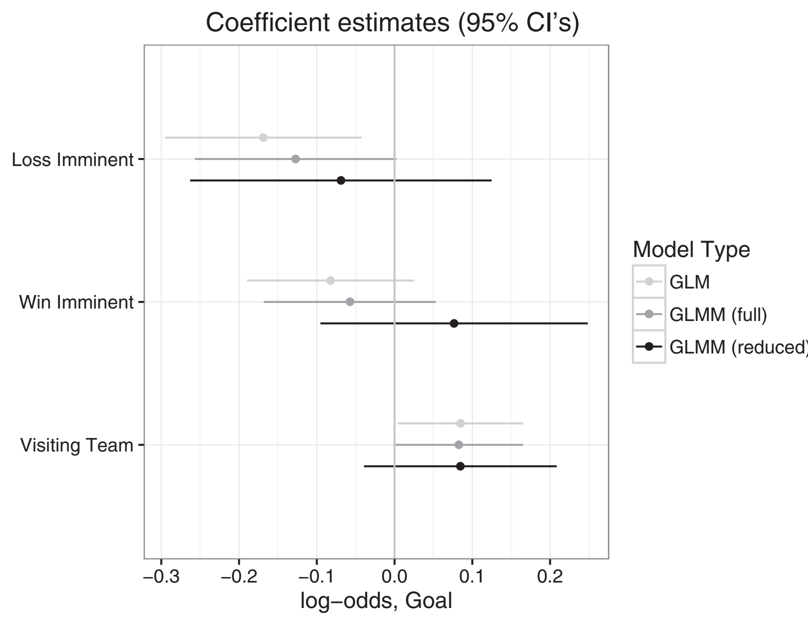

Figure 5. Coefficient estimates from mixed models of shoot-out outcomes, with 95% confidence intervals.

small. The posterior mean for the standard deviation of shooters ($\hat{\tau}_j^{-1} = 0.222$) is higher than that of goalies ($\hat{\tau}_k^{-1} = 0.127$). Although these scales are different than in the mixed models in Models (3) and (4), there is agreement that there is a larger variation among shooters than among goalies. The Bayesian model does not yield any goalies that have 99% credible intervals that do not include zero. Petteri Nummelin, Jakob Silfverberg, Victor Kozlov, Frans Nielsen and TJ Oshie rank as the top five shooters in the posterior summaries, as judged by using the lower bound of each player's credible interval among players whose intervals do not include zero.

Simulations

Resampling the shoot-out under randomness

Given that coaches are responsible for which shooters get the most opportunities, it is reasonable to expect that teams use past shooter performance in order to dictate strategy. Under this hypothesis, the best shooters would continue to get more attempts while shooters that struggle, either initially or eventually, would be passed over. We test this hypothesis as follows.

Let $SH_j(x)$ be the cumulative shooting percentage of shooter j after shot x, such that

$$SH_j(x) = \frac{\sum_{i=1}^{x} Goal_{ij}}{x}, \tag{5}$$

where $Goal_{ij}$ is an indicator for whether or not shooter j scored on shot i. Our interest lies in comparing how well the observed pattern of shooter deployment fits the expectations given a more random assignment of shooters to attempts.

To identify what $SH_j(x)$ would look like if gaps in shooter talent remained as wide as currently estimated but shooter allocation was independent of past success, we simulate under the following conditions. First, the overall league mean of each skater's goal percentage is taken to be the overall mean (33.25%, log-odds −0.695). Second, using $\hat{\tau}_j$, we assume that each shooter's random intercept comes from the $N(-0.695, \hat{\tau}_j = 0.191)$ distribution. We simulate intercepts for each shooter, transforming to get simulated probabilities. Each actual shooter (with known n_j total attempts) is assigned a simulated probability.

To compare the observed and simulated $SH_j(x)$, we use spaghetti plots (Figure 6). The left panel of Figure 6 shows observed $SH_j(x)$ for all j, while the right panel shows one example of a simulated $SH_j(x)$ for all j. In each panel, the grey line reflects each player's cumulative shoot-out percentage over time, while the black dot reflects the player's eventual percentage after all attempts are complete. Points are jittered to account for overlapping probabilities.

There are several differences between the observed and simulated tracks. In the simulated panel, far more players with high career percentages do not have additional opportunities. Related, in the simulated tracks panel, more players with a relatively poor performance are given additional opportunities. Meanwhile, the larger number of black dots in the bottom left of the observed panel suggests that in practice, most of these relatively poor shooters are no longer awarded opportunities.

Finally, the best players in the observed tracks panel have higher goal rates than in the simulated tracks, which suggests a possible skewed right distribution of the shooter-specific random intercepts. Most shooters converge around the overall mean of 33%; however, the poor ones are no longer given opportunities. Relative to what we would expect due to chance, the best shooters continue to outperform expectations.

Shooter and goalie value added

We estimate the relative importance of shooters and goalies with respect to team success in the shoot-out under two

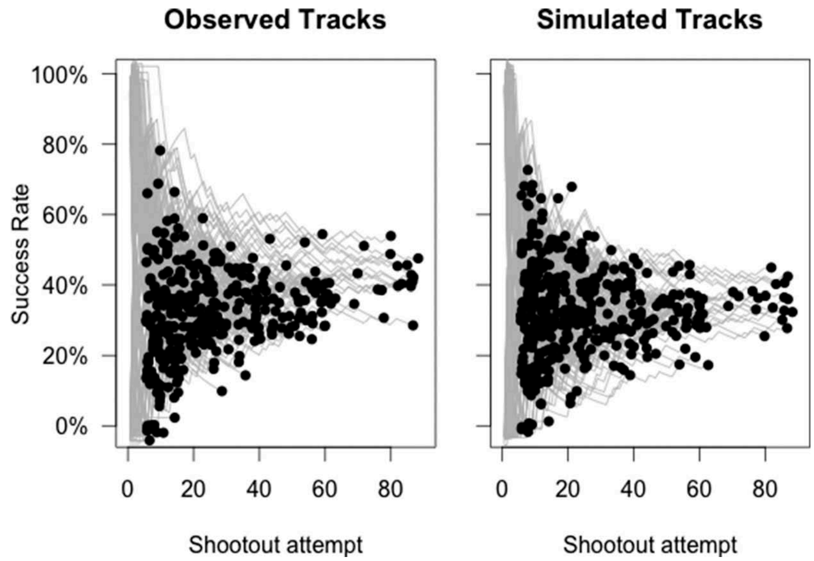

Figure 6. Observed and simulated shoot-out percentage tracks, given identical sample size of attempts.

assumptions, $A1$ and $A2$. Under $A1$, we estimate the net impact of adding one of the league's best shoot-out players (goal percentage, 51.7%) to a league average team, in terms of additional points gained per season towards seasonal standings. Under $A2$, we estimate the net impact of the league's best goalie (goal percentage, 26.3%) playing for a team comprised of league-average shooters.

Against league average teams, teams under $A1$ and $A2$ would win roughly 56% and 60% of shoot-outs, respectively. The minimum number of yearly shoot-outs a team played between the 2005–06 and 2014–15 seasons was 6, the mean 11.2, median 12 and the maximum was 21. Assuming that the number of games in a season that teams reach the shoot-out follows a Poisson distribution with parameter 11.2, we simulated 1000 team-seasons, comparing $A1$ and $A2$ to the performance of league average teams in terms of expected points added. Implicit in these simulations is that there is no association between a team's shoot-out ability and the frequency of shoot-outs they reach in a given season. This is reasonable; since the 2005–06 season, the correlation between a team's yearly shoot-out win percentage and its number of total shoot-outs is essentially zero (0.03).

Table 2 shows the mean, median, 2.5th and 97.5th quantiles of the simulated points added under $A1$–$A2$. Teams with the league's best shooter expect an additional 0.67 points towards the seasonal standings, on average, although in 47% of simulated seasons such a team would do worse than or the same as a league–average shoot-out team. Under $A2$, teams with the top goalie and league–average shooters do better than a league–average shoot-out team in 60% of seasons, picking up an average of 1.14 points per season.

Table 2. Net points towards season standings added.

Assumption	Mean points added	Median (2.5th–97.5th percentiles)
$A1$: add best shooter	0.67	1 (−4, 5)
$A2$: add best goalie	1.14	1 (−4, 6)

Using 1000 simulated seasons

Given that the relative value of an additional point in the standings for an NHL team has been valued at roughly a million American dollars (Patrick, 2014), it is reasonable to argue that, on shoot-out performance alone, Nielsen, Toews and Oshie have been worth roughly $670,000 in expectation, with Lundvquist and Fleury worth just over a million.

Incidentally, the expected shoot-out win percentages for teams with top shooters or goalies match empirical evidence. The New York Islanders, Chicago Blackhawks and the St. Louis Blues have won 54.5% of shoot-outs since the 2007–2008 season, the first in which shoot-out stars Nielsen, Toews and Oshie joined their respective teams. The Pittsburgh Penguins and the New York Rangers, featuring the league's best goalies in Fleury and Lundvquist since the 2005–06 season, have won 61.4% of shoot-outs over this time span.

Discussion & conclusion

The primary aims of this manuscript were to use graphical analyses, regression modelling and resampling to identify predictors of shot-level outcomes and the underlying randomness in shooter allocation and production in the NHL shoot-out. In contrast to much of the current literature, we present evidence that NHL shoot-outs are not an entirely random outcome, with the non-randomness coming in a couple of subtle forms.

First, although current literature on player skill in the shoot-out has argued that shoot-out results are what one would expect due to chance alone, we identify that there is significant non-random variability in player performance, at least among shooters. This is identified both using funnel plots and in regression modelling. Interestingly, while there is a larger variation in talent among shooters than goalies, it is more difficult for shooters to impact team performance as they, in all likelihood, only shoot once per contest. Given their performances over the past decade on shoot-outs alone, we find that shooters have been worth about twothirds of a point per season, and goalies worth about a point per season, in expectation, although in many seasons it is difficult to distinguish these differences from random variation.

Second, using simulations, we identify a selection bias with respect to shooter allocation; those with past success are repeatedly given opportunities, while those who miss on successive attempts are less likely to be chosen again. Related, we find that the best shooters have been used by their coaches in the first and second rounds of the shoot-out.

In addition to looking at claims of randomness, our regression model results also give information regarding correlates of player success. Although McEwan et al. (2012) identified that player performance varies under pressure, we find no such evidence after accounting for individual shooter and goalie talent. One plausible explanation for this difference is the round deployment of the best shooters, with the best ones going early in shoot-outs, which could unnecessarily link poor performance from the less talented shooters with performance under pressure. This claim is supported in the comparison of Models (3) and (4). Likely because there are several shooters without sufficient numbers of attempts in each round, Model (3) is unable to distinguish between shooter talent and the effect of pressure, and results suggest a link between pressure and performance. Model (4), however, reduces the need for this extrapolation; when including only shooters with enough attempts in each round, there appears to be no association between performance and pressure.

Finally, despite evidence that individual talent in the shoot-out is repeatable, it is difficult to distinguish team performance from season to the next from complete noise. From the perspective of team decision-makers, understanding the role that chance plays in shoot-out outcomes, and thus in the allocation of points in the standings, could be used to better interpret team performance.

Altogether, evidence suggests that there are few, if any, predictors of shoot-out success. Perhaps this should not come as a surprise, given that shoot-outs are a resistive exchange between shooter and goalie, with resulting variability reflecting their dynamic interactions. There does appear to be a moderate amount of within-shooter and a small amount of within-goalie variability in talent. Given the league's recent resistance to the shoot-out, noteworthy changes were made during the summer prior to the 2015–16 season. In place of a 5-min overtime with each team using four skaters apiece, teams are now playing with three skaters apiece. In anticipation of such a system, Pettigrew (2015) used historical scoring rates to estimate that the fraction of overtimes reaching a shoot-out would drop from roughly 60–43% with the implemented change. Assuming teams play overtime at the same frequency as in past seasons, we would expect between three and four fewer shoot-outs per team in each season. For team officials, this would result in a corresponding drop to the valuation of player performances mentioned earlier. Of course, if the frequency of overtime games continues to increase, as it has since the 2004–05 lockout (Lopez, 2015), the drop off may not be as severe.

While research in professional soccer has identified evidence of players choking under pressure (Jordet, 2009), no such evidence is found in hockey. One hypothesis for this difference is that in soccer, shooters control most of the shoot-out's outcome, as by and large, they can score with accurate kicks. In hockey, meanwhile, goalies control as much of the outcome, if not more, than the shooter. Relatedly, the overall fraction of goals in soccer lies around 80%, relative to the lower 33% goal rate in hockey. So, whereas more pressure may fall on shooters in soccer, perhaps shooters and goalies feel less pressure to succeed on any given attempt in hockey.

Disclosure statement

No potential conflict of interest was reported by the authors.

References

Albert, J. H., & Chib, S. (1993). Bayesian analysis of binary and polychotomous response data. *Journal of the American Statistical Association*, *88*, 669–679. doi:10.1080/01621459.1993.10476321

Apesteguia, J., & Palacios-Huerta, I. (2010). Psychological pressure in competitive environments: Evidence from a randomized natural experiment. *American Economic Review*, *100*, 2548–2564. doi:10.1257/aer.100.5.2548

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). lme4: Linear mixed-effects models using Eigen and S4 [Computer software manual]. Retrieved from http://CRAN.R-project.org/package=lme4 (R package version 1.1–8).

Brough, J. (2014). *Burke calls the shootout a circus stunt*. Retrieved December 31, 2015, from http://nhl.nbcsports.com/2014/02/28/burke-calls-the-shootout-a-circus-stunt/

Desjardins, G. (2009a, March). *Shootout length: Model vs actual*. Retrieved September 4, 2015, from http://behindthenet.ca/blog/2009/03/shootout-length-model-vs-actual.html

Desjardins, G. (2009b, October). *Shootouts: Does past performance mean anything?* Retrieved September 4, 2015, from http://www.arcticicehockey.com/2009/10/19/1089029/shootouts-does-past-performance

Desjardins, G. (2009c, October). *Shootouts: goaltender true talent*. Retrieved September 4, 2015, from http://www.arcticicehockey.com/2009/10/20/1089388/shootouts-goaltender-true-talent

Eric Tulsky (aka Eric T.). (2009). *Is the shootout really a crapshoot?* Retrieved June 15, 2015, from http://www.broadstreethockey.com/2012/2/13/2789709/yes-it-is

Feschuk, D. (2014, March). *No end in sight to the NHLs gimmick*. Retrieved September 4, 2015, from http://www.thestar.com/sports/hockey/2014/03/12/no_end_in_sight_to_nhls_gameending_gimmick_the_shootout_feschuk.html

Gelman, A. (2006). Prior distributions for variance parameters in hierarchical models. *Bayesian Analysis*, *1*(3), 515–533. doi:10.1214/06-BA117A

Gelman, A., Pasarica, C., & Dodhia, R. (2002). Let's practice what we preach: Turning tables into graphs. *The American Statistician*, *56*(2), 121–130. doi:10.1198/000313002317572790

Greenlees, I., Leyland, A., Thelwell, R., & Filby, W. (2008). Soccer penalty takers' uniform colour and pre-penalty kick gaze affect the impressions formed of them by opposing goalkeepers. *Journal of Sports Sciences*, *26*(6), 569–576. doi:10.1080/02640410701744446

Hurley, W. J. (2005). Overtime or shootout: Deciding ties in hockey. In J. Albert, J. Bennett, & J. J. Cochran (Eds.), *Anthology of statistics in sports*. ASA-SIAM Series on Statistics and Applied Probability, SIAM. Philadelphia, PA: SIAM.

Hurley, W. J. (2008). On the length of NHL shootouts. *Journal of Modern Applied Statistical Methods*, *7*. Retrieved from http://digitalcommons.wayne.edu/jmasm/vol7/iss1/27

Jones, M. (2013). The home advantage in individual sports: An augmented review. *Psychology of Sport and Exercise*, *14*, 397–404. doi:10.1016/j.psychsport.2013.01.002

Jordet, G. (2009). Why do English players fail in soccer penalty shootouts? A study of team status, self-regulation, and choking under pressure. *Journal of Sports Sciences*, *27*(2), 97–106. doi:10.1080/02640410802509144

Jordet, G., Hartman, E., Visscher, C., & Lemmink, K. A. P. M. (2007). Kicks from the penalty mark in soccer: The roles of stress, skill, and fatigue for kick outcomes. *Journal of Sports Sciences*, *25*(2), 121–129. doi:10.1080/02640410600624020

Leeds, M., & Von Allmen, P. (2004). *The economics of sports*. New York, NY: Pearson.

Lopez, M. (2015). Inefficiencies in the National Hockey League points system and the teams that take advantage. *Journal of Sports Economics*, *16*(4), 410–424. doi:10.1177/1527002513486654

McEwan, D., Ginis, K. A. M., & Bray, S. R. (2012). "With the game on his stick": The home (dis) advantage in national hockey league shootouts. *Psychology of Sport and Exercise*, *13*(5), 578–581. doi:10.1016/j.psychsport.2012.03.007

McGarry, T., & Franks, I. M. (2000). On winning the penalty shoot-out in soccer. *Journal of Sports Sciences*, *18*(6), 401–409. doi:10.1080/02640410050074331

Patrick, D. (2014). *How much does a standings point cost in the NHL*. Retrieved December 31, 2015, from http://www.fearthefin.com/2014/8/22/6055583/how-much-does-a-standings-point-cost-in-the-nhl

Peng, R. D. (2011). Reproducible research in computational science. *Science (New York, NY)*, *334*(6060), 1226–1227. doi:10.1126/science.1213847

Pettigrew, S. (2015). *How the 3 on 3 overtime rules would cut down NHL shootouts*. Retrieved December 31, 2015, from http://regressing.deadspin.com/how-those-3-on-3-overtime-rules-would-cut-down-nhl-shoo-1692120143

Pollard, R., & Reep, C. (1997). Measuring the effectiveness of playing strategies at soccer. *Journal of the Royal Statistical Society: Series D (The Statistician)*, *46*(4), 541–550. doi:10.1111/1467-9884.00108

R Core Team. (2015). *R: A language and environment for statistical computing [Computer software manual]*. Vienna, Austria. Retrieved from http://www.R-project.org/

Schuckers, M. (2009). *NHL shootout is a crapshoot*. Retrieved June 15, 2015, from http://www.arcticicehockey.com/2010/12/29/1901842/nhl-shootout-is-a-crapshoot

Schuckers, M., & Nelson, Z. (2013, January). *Impact of going first on winning an NHL shootout*. Retrieved from http://statsportsconsulting.com/main/wp-content/uploads/ShootoutFirst.pdf

Seppa, T. (2009, October). *Inside the VUKOTA projections: Shootouts*. Retrieved September 4, 2015, from http://www.hockeyprospectus.com/puck/article.php?articleid=305

Spiegelhalter, D. J. (2005). Funnel plots for comparing institutional performance. *Statistics in Medicine*, *24*(8), 1185–1202. doi:10.1002/(ISSN)1097-0258

Stubits, B. (2014, June). *NHL revenue to hit 3:7b; caplikelytoexceed70 million*. Accessed September 4, 2015, from http://www.cbssports.com/nhl/eye-on-hockey/24584103/report-nhl-revenue-to-hit-37-billion-cap-likely-to-exceed-70-million,note=

Wood, G., Jordet, G., & Wilson, M. R. (2015). On winning the "lottery": Psychological preparation for football penalty shoot-outs. *Journal of Sports Sciences*, *33*(17), 1758–1765. doi:10.1080/02640414.2015.1012103