DE GRUYTER Journal of Quantitative Analysis of Sports

# A Puck in the Net Beats Four Men in the Box



| | |
|---|---|
| Journal: | *Journal of Quantitative Analysis of Sports* |
| Manuscript ID: | DGJQAS.2014.0018 |
| Manuscript Type: | Research Article |
| Classifications: | |
| Keywords: | Hockey, Poisson, Markov, Win probability added, National Hockey League |



| | |
|---|---|
| Abstract: | We extend the classic Poisson model of hockey based on score differential and time remaining in the game to include the effect of penalties, and derive the associated Markov win probability model given the goal/manpower differential state at any point in a hockey game. Given data from the 2008/9-2011/12 National Hockey League seasons (a total of 4,920 games) reporting second-by-second goal and manpower differentials (which results in roughly 17.7 million observations), we estimate the state dependent transition rates and win probabilities. The data and model reveal that even after controlling for the home edge afforded by visiting teams being penalized more frequently than home teams, the goal scoring rate for the home team is higher than for visiting teams at most equivalent manpower differential levels. The results also show that, for any fixed duration of time left in the game, the probability that a team leading by x goals playing two men down wins is higher than the probability that a team leading by x-1 goals playing two men up wins, explaining the title of the paper. We demonstrate that this is a property of NHL hockey data as opposed to an artifact of the model. We use the model to develop a new win probability added metric for evaluating individual players based on their incremental contribution to the probability of winning and illustrate its use and conservation properties. Statistical issues that arise when comparing observed and modeled win probabilities are examined in an appendix. |

# A Puck in the Net Beats Four Men in the Box

February 2014

**Abstract**

We extend the classic Poisson model of hockey based on score differential and time remaining in the game to include the effect of penalties, and derive the associated Markov win probability model given the goal/manpower differential state at any point in a hockey game. Given data from the 2008/9-2011/12 National Hockey League seasons (a total of 4,920 games) reporting second-by-second goal and manpower differentials (which results in roughly 17.7 million observations), we estimate the state dependent transition rates and win probabilities. The data and model reveal that even after controlling for the home edge afforded by visiting teams being penalized more frequently than home teams, the goal scoring rate for the home team is higher than for visiting teams at most equivalent manpower differential levels. The results also show that, for any fixed duration of time left in the game, the probability that a team leading by $x$ goals playing two men down wins is higher than the probability that a team leading by $x-1$ goals playing two men up wins, explaining the title of the paper. We demonstrate that this is a property of NHL hockey data as opposed to an artifact of the model. We use the model to develop a new win probability added metric for evaluating individual players based on their incremental contribution to the probability of winning and illustrate its use and conservation properties. Statistical issues that arise when comparing observed and modeled win probabilities are examined in an appendix.

# 1 Introduction

The game of hockey differs from other team sports such as baseball, basketball, football and soccer in how teams are disciplined for player or team infractions, or *penalties*. Typically, when a penalty is called in hockey, the offending player is removed from the ice at the next stoppage in play, which forces the penalized team to play shorthanded until the penalty expires or the opposing team scores, whichever happens first. Since goal-scoring rates are higher for teams skating on a "power play" and lower for teams engaged in "penalty killing," penalties are often crucial in determining the outcome of a hockey game. In this article, we develop a new model of the probability of winning a hockey game that explicitly accounts for the manpower differentials caused by penalties during a game. Calibrated with over 17.7 million seconds of hockey data describing every regular season hockey game in the 2008-09 through 2011-12 seasons, we show how this model can be used to create a real-time win probability scorecard that in turn can be used to evaluate a hockey player's individual contribution to the probability of winning (win probability added).

We are by no means the first to model hockey, and our research owes much to earlier work. Mullet (1977) established the usefulness of the Poisson distribution in describing the number of goals scored during hockey games using data from the 1973-74 National Hockey League (NHL) season. Buttrey, Washburn, and Price (2011) used a Poisson model to analyze per-minute goal scoring rates across different man-advantage scenarios using data from the 2008-09 season. Using in-game Poisson models, several authors examined the consequences and optimal timing of "pulling the goalie," a unique hockey strategy whereby an extra skater is substituted for the goaltender, typically employed by the trailing team towards the end of a game. These include Morrison (1976), Morrison and Wheat (1986), Erkut (1987), Nydick and Weiss (1989), Washburn (1991), Berry (2000), Zaman (2001), and Beaudoin and Swartz (2010), which develops a simulation program that considers various man-advantage situations and team-specific parameters. Collectively, these papers overwhelmingly suggest that teams wait too long to pull their goaltender.

Our research builds most directly on Washburn (1991), who used the Poisson process to develop a state-space model of the probability that a team wins a hockey game given the goal differential and time left in the game.[1] Washburn (1991) did not consider the effect of penalties on goal scoring, but Buttrey, Washburn and Price (2011) and Beaudoin and Swartz (2010) did. In those papers, scoring rates were estimated for specific man-advantage situations (e.g. 6-on-5 or 5-on-4), whereas in our work we focus only on manpower differential (thus both 6-on-5 and 5-on-4 would be modeled as a manpower differential of +1) for purposes of creating a tractable state-space model that produces win probabilities given the goal and manpower

[1]Stern (1994) develops similarly-spirited models for basketball based on Brownian motion that, while appropriate for basketball, is not well-suited for hockey where scoring is rare and the natural stochastic process is Poisson.

differential at any point in a game.

The paper unfolds as follows: in the next section we briefly review the Poisson model of hockey where the state-space depends only upon goal differential at any point in the game. We then develop a new Markov model for the probability that the home team wins a hockey game given manpower differential in addition to goal differential at any point in the game. In Section 4 we discuss our dataset, which consists of the home-minus-away goal and manpower differential at every second of every regular season hockey game over four NHL seasons, estimate the required state transition rates, and discuss some properties of home ice advantage that emerge from the data. Section 5 presents our numerical win probability results for the model of Section 4, and uses the model to explore the earlier analysis of Burke (2009) regarding the relative boost in win probability provided by gaining a man advantage via a penalty relative to scoring a goal. We also show that while penalties matter, one would never trade a goal for a four-man swing in manpower differential, that is, a puck in the net beats four men in the box. In Section 6, we demonstrate how to develop a real-time win probability scoreboard for keeping track of a hockey game; as opposed to changing score only when goals occur, our win probability scoreboard changes continuously over the duration of a hockey game in addition to discrete jumps that occur when goals are scored and penalties are awarded or expire. Section 7 applies this scoreboard to evaluate individual players by generalizing the +/- statistic to individual win probability added, provides a numerical example, and shows how these individual measures cumulate to team wins above average over the course of a season. We conclude in Section 8, while the importance of model-implied variability in observed win probabilities for assessing the fit of the Markov model to the data is detailed in the appendix, where we derive explicit formulas for the variance of observed win probabilities based on the model.

## 2 The Poisson Model of Hockey

We begin with a quick review of the Poisson model of hockey (see Washburn (1991) for an application to pulling the goalie). Let $\lambda$ ($\mu$) denote the expected number of goals scored by the home (away) team per unit time in a hockey game. Let $w(x,t)$ denote the conditional probability that the home team wins the hockey game, given that with $t$ time units left in regulation, the home team leads the visitors by $x$ goals (if $x$ is negative, the home team trails). The Poisson model of hockey states that the home and away teams score in accord with independent Poisson processes with scoring rates $\lambda$ and $\mu$ respectively. Consequently, given that the current home minus away goal differential equals $x$, in any time slice of duration $\Delta t$, three events are possible (Figure 1): (i) the home team scores, in which case the goal differential increases from $x$ to $x+1$ (this event occurs with probability $\lambda\Delta t$); (ii) the away team scores, in which case the goal differential decreases from $x$ to $x-1$ (this even occurs with probability $\mu\Delta t$); and (iii) no team scores in which case the goal differential remains at $x$ (this event occurs with probability $1-(\lambda+\mu)\Delta t$).

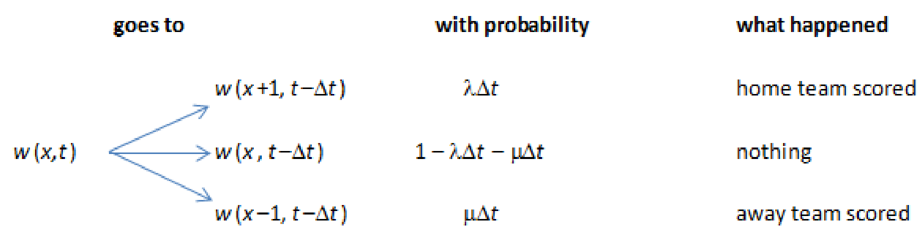


Figure 1: The Poisson Model of Hockey

These dynamics lead to equation (1) below which applies to regulation time of a hockey game. The model is closed by equation (2), which indicates that the home team wins with certainty if it has outscored the away team by the end of regulation time at $t=0$ (the away team wins with certainty if it has scored more goals), and for simplicity assigns a winning probability of 1/2 if the game is tied at the end of regulation time (later we will explicitly model overtime). As an illustrative example, consider the four NHL regular seasons spanning 2008-2012 (1,230 games per season; 4,920 games total). Over all of these games, the home team averaged 2.75 goals per game while the visiting team averaged 2.47 goals per game. Using these averages to estimate $\lambda$ and $\mu$ (recall that there are 60 minutes in regulation time), equations (1)-(2) predict that the home team wins 54.7% of the time (that is, $w$(0,60 minutes) = 0.547). In fact, over the 4,920 games considered, the home team won 2,702 or 54.9% of the games.

$$\frac{dw(x,t)}{dt} = \lambda w(x+1,t) + \mu w(x-1,t) - (\lambda+\mu)w(x,t) \tag{1}$$

$$w(x,0) = \begin{cases} 1 & x > 0 \\ \frac{1}{2} \quad \text{if} & x = 0 \\ 0 & x < 0 \end{cases} \tag{2}$$

While this is impressive considering that no win/loss data were taken into account in calibrating this model, there is more to a hockey game than simply scoring goals. As discussed in the introduction, one feature distinguishing hockey from other team sports is the imposition of penalties for various infractions that deliberately create unbalanced manpower situations. Since scoring rates are not constant across manpower advantage scenarios, an assertion with strong intuitive and empirical support, it stands to reason that the probability of winning a hockey game depends on manpower differential in addition to goal differential.

## 3 A Markov Model Incorporating Manpower Differential

We now proceed to incorporate manpower differential into the state space to model the conditional probability that the home team wins a hockey game, given that the home team leads the visiting team by $x$ goals while enjoying a manpower differential of $y$ players with $t$ time units remaining in regulation. For example, if the home team is trailing by two goals while playing with a man advantage, the $(x, y)$ state of the game is $(-2, 1)$. Following NHL rules, we only consider manpower differentials $y = -2, -1, 0, 1$, or 2, while for computational convenience we apply a five goal mercy rule that artificially terminates a game in favor of the home team if $x = 5$ (the visitors win if $x = -5$). The instantaneous state transition rate from a goal/manpower differential state $(x, y)$ to a different state $(x', y')$ is denoted by $\lambda_{xy}^{x'y'}$. These transition rates can, for examples, reflect home (away) team scoring when $x' = x + 1$ $(x - 1)$, or an increment (decrement) in manpower differential as occurs when a penalty against the visiting team occurs (expires) in which case $y' = y + 1$ $(y - 1)$. Denoting the state-dependent home team win probability by $w(x, y, t)$ and measuring time in seconds (so regulation time equals 3,600 seconds), the equations of the model are:



$$\frac{dw(x,y,t)}{dt} = \sum_{(x',y') \neq (x,y)} \lambda_{xy}^{x'y'} w(x', y', t) - \left( \sum_{(x',y') \neq (x,y)} \lambda_{xy}^{x'y'} \right) w(x, y, t) \qquad (3)$$

$$x = -4, ..., 4;\ y = -2, ..., 2;\ 0 < t \leq 3600$$

$$w(x,y,0) = \begin{cases} 1 & x > 0 \\ w(0,y,0) & x = 0 \\ 0 & x < 0 \end{cases} \quad x = -5,...,5;\ y = -2,...,2 \tag{4}$$

$$w(5,y,t) = 1; w(-5,y,t) = 0;\ y = -2,...,2;\ 0 < t \leq 3600 \tag{5}$$

$$\frac{dw(0,y,t)}{dt} = \sum_{(x',y')\neq(0,y)} \lambda_{0y}^{x'y'} w(x',y',t) \quad - \quad \left( \sum_{(x',y')\neq(0,y)} \lambda_{0y}^{x'y'} \right) w(0,y,t) \tag{6}$$
$$y = -2,...,2;\ -300 < t \leq 0$$

$$w(1,y,t) = 1; w(-1,y,t) = 0;\ y = -2,...,2;\ -300 < t \leq 0 \tag{7}$$

$$w(0,y,-300) = \frac{1}{2};\ y = -2,...,2 \tag{8}$$

Equation (3) describes the evolution of the win probability during regulation time. Equation (4) awards a certain win to the home (away) team if, at the end of regulation time, the home (away) team has scored more goals in the game. If the game is tied at the end of regulation however, equation (4) assigns a win probability that depends upon the resolution of overtime play as modeled via equation (6). Equation (5) assigns a win with certainty to the home or away team if either achieves a five goal advantage, which is taken to be insurmountable. Equation (6) mimics equation (3) in describing the win probability evolution during overtime play, except that the goal differential is restricted to zero as long as overtime ensues. Of course, if either the home or away team score during sudden-death overtime, the game ends with victory for the scoring team; this is enforced by equation (7). Overtime cannot last more than five minutes (300 seconds) after which a penalty shootout ensues; equation (8) assigns a win probability of 1/2 to both teams in such an event.

# 4 Data, Parameter Estimation, and Home Ice Advantage

## 4.1 The Data

Our initial data were drawn from the play-by-play reports that are publicly available at nhl.com. In addition to the identity of the home and away teams, these reports record and time-stamp many different events that take place during games, including goals scored and penalties. For each goal and penalty, the reports contain associated information relating to the identity of the team and the number of team-goals scored prior to the occurrence of the event. In the case of penalty events, the reports contain associated information relating to the penalty duration (e.g., 2 minutes). Based on the rules of the games that determine the expiration of penalty events from either elapsed time or opposing team goals scored, an algorithm was developed to identify the score and manpower differential states for each game-second. Our final data set consists of over 17.7 million season-game-second-level specific score and manpower differential states, relative to the home team, during games played from the 2007-08 through 2011-12 seasons.[2]

## 4.2 Parameter Estimation

Estimating the state-to-state transition rates from the data described above is straightforward in principle. For any two states $s = (x, y)$ and $s' = (x', y') \neq s$, define $n(s, s')$ as the number of transitions observed in the data from state $s$ to state $s'$, and let $\tau(s)$ denote the total time spent in state $s$ over the entire data set. For example, if $s = (0, 0)$, then aggregating over all 4,920 hockey games, $\tau(s)$ is the total time spent by teams playing at even strength while the score is tied. With these definitions, the maximum likelihood estimate for the transition rate $\lambda_s^{s'}$ is equal to

$$\lambda_s^{s'} = \frac{n(s, s')}{\tau(s)}. \tag{9}$$

However, with nine goal differential ($x = -4, ..., 4$) and five manpower differential ($y = -2, ..., 2$) possibilities (see equation (3)), there are 45 possible states and thus $45 \times 44 = 1,980$ conceivable state-to-state transition rates. Most such transitions are impossible. For example, the goal differen-

[2] 17,712,000 seconds = 60 seconds/minute × 60 minutes/game × 1230 games/season × 4 seasons

tial cannot change by more than one in absolute value in any single state transition. We therefore decided to use the much smaller set of parameters that results from removing the dependence of state transition rates upon initial goal differential, and only allowing transitions that would change goal differential by at most one. An implication of this assumption is that while goal scoring rates depend upon manpower differential, they do not depend upon goal differential; this assumption generalizes the simple Poisson model where goal scoring rates are completely independent of both goal and manpower differential. The estimated transition rates per game appear in Table 1.

| Change in Goal Differential | x' - x | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | -1 | -1 | -1 | -1 | -1 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Final Manpower Differential | y' | 2 | 1 | 0 | -1 | -2 | 2 | 1 | 0 | -1 | -2 | 2 | 1 | 0 | -1 | -2 |
| Initial | 2 | 0.00 | 11.92 | 0.37 | 0.00 | 0.00 | --- | 35.24 | 3.29 | 0.00 | 0.00 | 0.48 | 0.04 | 0.00 | 0.00 | 0.00 |
| Manpower | 1 | 0.00 | 0.16 | 5.99 | 0.00 | 0.00 | 2.45 | --- | 28.77 | 0.08 | 0.01 | 0.00 | 0.86 | 0.03 | 0.00 | 0.00 |
| Differential | 0 | 0.00 | 0.00 | 2.49 | 0.00 | 0.00 | 0.07 | 4.22 | --- | 3.92 | 0.06 | 0.00 | 0.00 | 2.25 | 0.00 | 0.00 |
| ($y$) | -1 | 0.00 | 0.00 | 0.01 | 0.97 | 0.00 | 0.01 | 0.09 | 28.98 | --- | 2.17 | 0.00 | 0.00 | 5.59 | 0.12 | 0.00 |
| | -2 | 0.00 | 0.00 | 0.00 | 0.00 | 0.65 | 0.00 | 0.04 | 3.41 | 33.65 | --- | 0.00 | 0.00 | 0.13 | 12.40 | 0.00 |

Table 1: State transition rates per 3600 seconds of hockey.

### 4.3 Home Ice Advantage

Note that since a transition with $x' - x = 1$ implies that the home team scored a goal, summing along the rows across the first five columns of Table 1 produces the home team scoring rates for each manpower differential; similarly summing along the rows across the last five columns (corresponding to $x' - x = -1$) yields the manpower differential dependent scoring rates for the visiting team. These scoring rates are displayed in Table 2, which makes clear how important manpower differential is in hockey.[3] For both the home and away teams, the scoring rates increase monotonically with manpower advantage; indeed there is about a 20-fold increase in the goal scoring rates as teams move from playing two men short to having a two man advantage.

| **Manpower State** | **Home Goals/Game** | **Away Goals/Game** |
|---|---|---|
| Home Up 2 | 12.28 | 0.51 |
| Home Up 1 | 6.15 | 0.89 |
| Even Strength | 2.49 | 2.26 |
| Home Down 1 | 0.99 | 5.71 |
| Home Down 2 | 0.65 | 12.53 |
| **Overall Average** | **2.75** | **2.47** |

Table 2: Goal scoring rates per 3600 seconds of hockey for different manpower differentials.

There is an additional observation of interest regarding home ice advantage that follows directly from these rates. That home teams enjoy an edge in baseball, basketball, and football is well known (Moskowitz and Werthem, 2011), and hockey is no exception; as discussed earlier, in our data the home team averages 2.75 goals per game while the away team averages 2.47. Moskowtiz and Werthem (2011) have argued that the main source of the home edge stems from referee bias in favor of the home team, perhaps in response to pressure from exuberant home team fans; in hockey, this would translate into referees calling more penalties against the visiting team than

[3] These rates are broadly consistent with those reported in Table 1 of Buttrey, Washburn and Price (2011) based on one NHL season that conditions on the actual number of skaters on ice (and not just their difference), but does not distinguish scoring rates for home versus away. See also Beaudoin and Swartz (2010).

the home team, and indeed this does occur: over the four years in our data, home teams were penalized on average 4.64 times per game while the visiting team averaged 4.99 penalties per game (see also Beaudoin and Swartz (2010)). However, Table 2 shows that the home team outscores the visiting team at four out of five manpower differential levels (the exception being a two man advantage). For example, at even strength, the home team scores 2.49 goals per 60 minutes of hockey while the away team scores only 2.26 goals per (regulation) game. Playing with a one man advantage, the home team scores 6.15 goals per 60 minutes of hockey; when the away team has a one man advantage, it averages 5.71 goals in 60 minutes. Playing two men down, the home team scores 0.65 goals per 60 minutes, which is low, yet higher than the 0.51 goals per 60 minutes the visiting team averages in 60 minutes playing two men down. Given these results, it seems that the home edge cannot be completely explained by referees penalizing visiting teams more often.

# 5 In-Game Win Probabilities Incorporating Manpower Differential

Using the estimated transition rates reported in Table 1, we numerically solved equations (3)-(8) to obtain $w(x, y, t)$, the home team win probability given the home-minus-away goal ($x$) and manpower ($y$) differentials with $t$ seconds remaining in the game. Win probabilities for goal differentials between $-2$ and $+2$ at all manpower differential levels are displayed in Figure 2 (goodness-of-fit details appear in the Appendix). The curves in this graph neatly divide into five different groups of five curves each, one group for each goal differential and five curves within each group corresponding to the five different manpower differentials. Focusing on the middle group first, at the start of a game, the score is tied, teams are at even strength, and the home team win probability equals 55%, consistent with the observed data. If the home team rapidly gains a two man advantage near the start of the game, this increases the win probability to 61%, while if the home team rapidly finds itself down two men, the win probability drops to 49%. By the end of regulation time (3,600 seconds), the home win probability ranges from 35% if the home team enters overtime playing two men down to 66% if the home team enjoys a two man advantage; this is the widest margin due to penalties that can occur at any time during a game for any fixed goal differential. Starting overtime at equal strength (meaning four skaters per team as per NHL rules as opposed to the usual five skaters), the home team win probability equals 51%. If there is no scoring during five minutes of overtime play, then the win probability equals exactly 50% via equation (8); this explains why the win probabilities converge to 1/2 for all manpower differentials in a tie game.

The other four groups of curve correspond to games where the home team has a two goal lead (highest set of five curves), a one goal lead (next highest set of five curves), a one goal deficit (second lowest set of five curves), and a two goal deficit (lowest set of five curves). All of these win probability curves converge to either one or zero depending upon whether the home team is leading or trailing when regulation time expires.

Figure 2 shows the impact of manpower differential on the probability of winning a hockey game. The reason, of course, is easily understood from Table 2: if the home team is playing with a one or two man advantage (and consequently the visitors face a one or two man deficit), the difference between home and away team scoring rates increases from 0.23 per 60 minutes of hockey to 5.24 and 11.77 per 60 minutes respectively. Scoring a goal shifts the goal differential by one, placing the home team in the next highest set of win probability curves. Conversely, if the home team faces a one or two man deficit, then the home minus away goal scoring rate declines from 0.23 per 60 minutes to $-4.72$ and $-11.88$ respectively; when the away team scores, the home team goal differential declines by one, shifting the home team in the next lowest set of win probability curves.

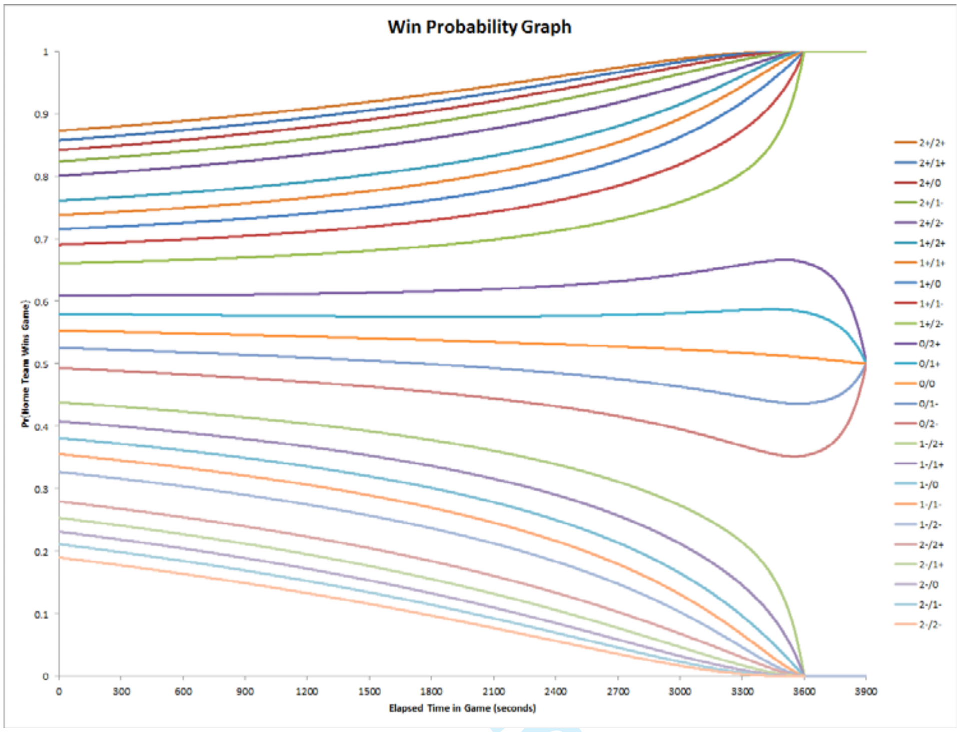


Figure 2: In-game NHL goal and manpower differential-dependent home team win probabilities based on transition rates of Table 1.

Since teams playing with a man advantage score on approximately 20% of power plays while teams skating shorthanded score on only 2% of such occasions, Brian Burke (2009) reasoned that gaining a man advantage should increase the probability of winning by a little less than 20% of the difference between the win probability with the current goal differential and the win probability at the next highest goal differential based on the Poisson model ignoring manpower differential (see equations (1)-(2)). We can check Burke's intuition using Figure 2 by calculating $((w(x, y+1, t) - w(x, y, t))/(w(x+1, y, t) - w(x, y, t))$, which is the fraction of the win probability jump from scoring a goal at a fixed manpower level that is covered by the win probability jump from gaining a man advantage without scoring a goal, and comparing the result to 20%.

Figure 3 reports one such comparison, where we consider the impact of increasing manpower differential by one from even strength in games where the game is tied or the home team leads or trails by one goal. This figure supports Burke's observation over the first half of regulation play, in that after 30 minutes of hockey, the impact of gaining a man advantage accounts for between 15% (when the home team is trailing by a goal) and 19% (when the home team is up a goal) of the incremental win probability a goal would bring. However, the impact of penalties relative to goals changes dramatically later in the game. Gaining a man advantage while playing a goal up is worth about 66% of the increment in win probability that a second goal would bring near the end of regulation time. Going a man up with a goal lead late in regulation yields a very high probability of preserving the lead and winning the game that is equivalent to 66% of the lock on a win a second goal would provide. Conversely, a team trailing by a goal towards the end of regulation is still quite likely to lose if it gains a late one man advantage, which provides almost none of the advantage that a goal (and a trip to overtime) would provide. Burke's rule of thumb applies best to tie games, where winning a penalty provides 17% of the value of scoring a goal towards the beginning of a game, and declines only slightly to 15% by the end of regulation. However, in overtime the value of gaining a man advantage relative to a goal declines sharply, which is not a surprise given that a goal wins the game.

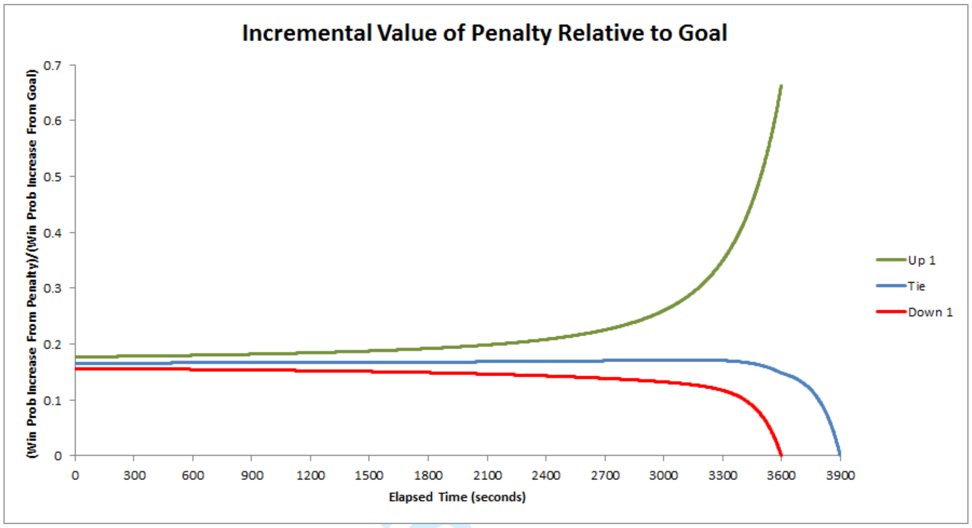


Figure 3: Percentage increase in win probability from winning a penalty versus scoring a goal.

## 5.1 A Puck in the Net Beats Four Men in the Box

Note that the win probability curves in Figure 2 *never cross*. This means that the probability of winning strictly increases with goal differential, without regard to manpower differential, at any point in the game. In particular, consider the win probability curve that results when the home team enjoys a two man advantage ($y = 2$) playing in a tie game ($x = 0$); this is the purple 0/2+ curve in Figure 2. Compare this to the win probability curve that results when the home team is playing two men down ($y = -2$) with a one goal lead ($x = 1$); this is the green 1+/2- curve. The green curve is always above the purple curve, which means that $w(1, -2, t) > w(0, 2, t)$ for all values of time $t$ remaining in the game. Consequently, the home team would always prefer to be in the $(x, y)$ state $(1, -2)$ as opposed to $(0, 2)$. Suppose it was possible to exchange state $(1, -2)$ for $(0, 2)$, that is, suppose the home team could give up a goal in exchange for a net gain of four skaters (achieved by retrieving two home team players from and sending two visiting players to the penalty box). The home team would, of course, never make such an exchange. That the curves in Figure 2 never cross implies that the result just described applies more generally: in Figure 2, $w(x, -2, t) > w(x-1, 2, t)$ for all values of $t$. For any duration of time left in the game, the probability that a team leading by $x$ goals playing two men down wins is higher than the probability that a team leading by $x-1$ goals playing two men up wins. Consider the win probability curve that results when the home team enjoys a two man advantage ($y = 2$) playing at a goal differential of $x - 1$, and compare this to the win probability curve that results when the home team is playing two men down ($y = -2$) at a goal differential of $x$. Figure 2 shows that *a puck in the net beats four men in the box*, that is, gaining a goal is *always* worth more than a favorable four person change in manpower differential. Given this result, it is generally a good idea for a team to take a penalty to save a goal!

This is a property of NHL hockey as opposed to an artifact of our hockey model. To see this, recall that Figure 2 is dependent upon the state-to-state transition rates estimated from $\approx 17.7$ million seconds of regular season NHL hockey that in turn produced the goal scoring rates reported in Table 2. Now imagine changing the state transition rates in the first row of Table 1 in such a manner that the home team scoring rate with a two man advantage increases by a factor of about 7 while holding all other transition (and hence scoring) rates constant. The resulting win probability curves from such distorted transition rates are shown in Figure 4.

Consider the dotted green curve that corresponds to the win probability when the home team has a two man advantage in a game they trail by one goal (1-/2+). This curve is now *higher* than the red win probability curve describing tie games where the home team is playing two men down (0/2-). In this fictional case, the home team would rather play two men up in a game they trail by one goal than two men down in a tie. The reason is that with such a high goal scoring rate with a two man advantage, the home team is quite likely to score which, upon releasing a visiting skater from the penalty box, would place them in a tie game with a one man advantage, which is much better than being in a tie game playing down two men. If the scoring rate with a two man advantage was increased even further, then the win probability curve when trailing by a goal with a two man advantage would approach the win probability curve for a tie game with a one man advantage for this same reason. Note that $w(x-1,2) > w(x,1)$ for all goal differentials $x$ in this example. By constructing the counterexamples of Figure 4, we have shown that the non-crossing win probability curves of Figure 2 are a property of NHL hockey as opposed to a model artifact.

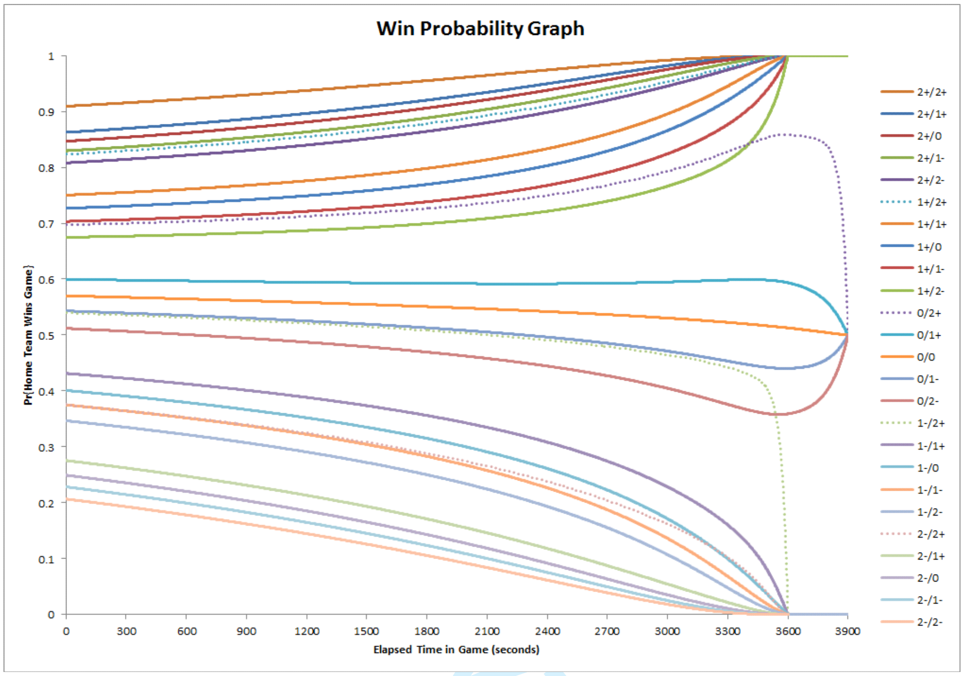


Figure 4: Hypothetical win probability when home goal scoring rate with two man advantage increases by a factor of seven.

## 6 Real-Time Win Probability Scoreboard

The win probability model reported in Figure 2 can be used to assess the in-game probability that the home team will win a hockey game based on the goal and manpower differential at any point in time during the game. To illustrate, Table 3 reports the official NHL game summary for the October 9, 2010 game between the visiting Dallas Stars and the home team New York Islanders. This summary contains all of the information necessary to deduce the goal and manpower differentials at any point in the hockey game providing one understands the rules of hockey. For example, the New York penalty that occurred at 8:49 of the first period gave Dallas a man advantage (so $y = -1$ since New York is the home team), while the power play Dallas goal scored at the 9:52 of the first period ended the penalty and returned the game to even strength ($y = 0$) but gave Dallas a one goal lead ($x = -1$).

Figure 5 shows the result of mapping all of the game state information inferred from Table 3 onto the win probability graph of Figure 2; we refer to the result as the real-time win probability scoreboard. The mapping simply selects the appropriate win probability curve corresponding to the state of the game at all points in time. Specifically, the underlying state-specific curves from Figure 2 are shown as dotted gray lines in Figure 5, while the solid blue curve highlights the win probability associated with the actual goal/manpower differential game state at all points during the game. Note how the Dallas goal scored at 9:52 of the first period (592 seconds into the game) manifests as a vertical drop from the curve corresponding to the home team playing down one man in a tie game to the curve corresponding to even strength while trailing by a goal. This corresponds to a drop in the win probability from 0.52 to 0.36, thus the combined effect of surrendering a goal and regaining a skater dropped New York's win probability by 16 percentage points. Note that while this example was constructed after the fact from Table 3, there is absolutely no reason why graphs like Figure 5 could not be constructed in real time using the publicly available live data feed from nhl.com.

VISITOR

HOME

**Game Summary**

5     4

Saturday, October 9, 2010
Attendance 13,351 at Nassau Coliseum
Start 7:16 EDT; End 9:55 EDT
Game 0015
Final

DALLAS STARS
Game 2 Away Game 2

NEW YORK ISLANDERS
Game 1 Home Game 1

SCORING SUMMARY

| G | Per | Time | Str | Team | Goal Scorer | Assist | Assist | DAL on Ice | NYI on Ice |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 9:52 | PP | DAL | 14 J.BENN(1) | 91 B.RICHARDS(2) | 29 S.OTT(1) | 6, 10, 14, 29, 32, 91 | 4, 24, 26, 39, 51 |
| 2 | 1 | 11:59 | EV | DAL | 3 S.ROBIDAS(1) | 91 B.RICHARDS(3) | 6 T.DALEY(1) | 3, 6, 18, 21, 32, 91 | 4, 24, 26, 39, 91 |
| 3 | 2 | 6:21 | PP | NYI | 20 J.WISNIEWSKI(1) | 93 D.WEIGHT(1) | 15 P.PARENTEAU(1) | 3, 14, 29, 32, 37 | 12, 15, 20, 26, 39, 93 |
| 4 | 2 | 13:45 | EV | DAL | 10 B.MORROW(2) | 63 M.RIBEIRO(2) | 2 N.GROSSMAN(1) | 2, 3, 10, 16, 32, 63 | 10, 15, 20, 26, 28, 39 |
| 5 | 2 | 15:24 | PP | NYI | 93 D.WEIGHT(1) | 15 P.PARENTEAU(2) | 20 J.WISNIEWSKI(1) | 3, 32, 37, 63 | 12, 15, 20, 26, 39, 93 |
| 6 | 3 | 7:12 | EV | DAL | 10 B.MORROW(3) | 63 M.RIBEIRO(3) | | 5, 10, 29, 32, 37, 63 | 16, 27, 39, 47, 57, 93 |
| 7 | 3 | 8:45 | EV | NYI | 57 B.COMEAU(1) | 24 R.MARTINEK(1) | 12 J.BAILEY(1) | 6, 13, 28, 29, 32, 91 | 4, 12, 24, 28, 39, 57 |
| 8 | 3 | 17:21 | PP | NYI | 26 M.MOULSON(1) | 51 F.NIELSEN(1) | 93 D.WEIGHT(2) | 2, 10, 28, 32, 63 | 7, 20, 26, 39, 51, 93 |
| 9 | SO | | | DAL | 63 M.RIBEIRO | | | | |

PENALTY SUMMARY

**DALLAS STARS**

| # | Per | Time | Player | PIM | Penalty |
|---|---|---|---|---|---|
| 1 | 1 | 1:53 | 29 S.OTT | 2 | Hooking |
| 2 | 1 | 3:44 | 18 J.NEAL | 2 | Tripping |
| 3 | 2 | 6:15 | 17 T.PETERSEN | 2 | Hooking |
| 4 | 2 | 14:06 | 2 N.GROSSMAN | 2 | Interference |
| 5 | 2 | 14:31 | 29 S.OTT | 2 | Tripping |
| 6 | 3 | 10:56 | 91 B.RICHARDS | 2 | Holding |
| 7 | 3 | 16:17 | 3 S.ROBIDAS | 2 | Delaying Game-Puck over glass |
| 8 | 3 | 17:53 | 3 S.ROBIDAS | 2 | Delaying Game-Puck over glass |
| TOT (PN-PIM) 8-16 | | | | | |
| Power Plays (Goals-Opp./PPTime) 1-7/11:12 | | | | | |

**NEW YORK ISLANDERS**

| # | Per | Time | Player | PIM | Penalty |
|---|---|---|---|---|---|
| 1 | 1 | 8:49 | 57 B.COMEAU | 2 | Boarding |
| 2 | 1 | 16:59 | 4 M.EATON | 2 | Interference |
| 3 | 2 | 3:51 | TEAM | 2 | Too many men/ice - bench |
| 4 | 2 | 10:46 | 93 D.WEIGHT | 2 | Hi-sticking |
| 5 | 2 | 19:40 | 20 J.WISNIEWSKI | 2 | Slashing |
| 6 | 3 | 10:47 | 16 J.SIM | 2 | Roughing |
| 7 | OT | 0:29 | 57 B.COMEAU | 2 | Cross checking |
| TOT (PN-PIM) 7-14 | | | | | |
| Power Plays (Goals-Opp./PPTime) 3-8/09:35 | | | | | |

Table 3: NHL game summary from http://www.nhl.com/scores/htmlreports/20102011/GS020015.H

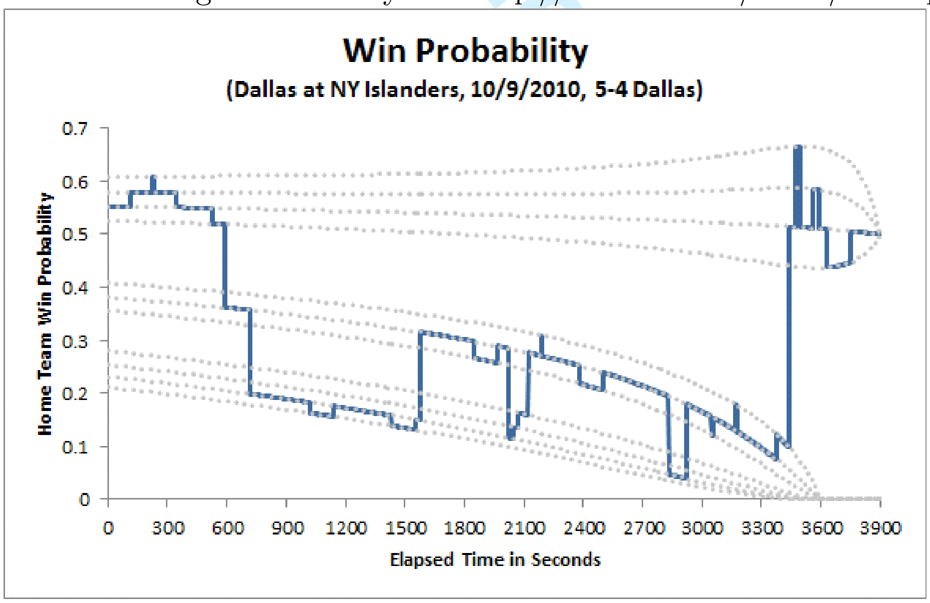


Figure 5: The win probability scorecard for Dallas at NY Islanders based on the events in Table 3.

It is important to point out what the win probabilities in Figure 5 actually mean, however. It would not be accurate to say that the figures graphed correspond to the actual probabilities that New York defeats Dallas, for the underlying model from Figure 2 was calibrated for all games over four seasons. However, while the win probabilities are themselves based on all games, the state transitions that give rise to the jumps in Figure 5 are the actual game transitions that result from the play of the two teams in the game. Thus, the model creates an alternative scoreboard. Rather than keep track in a game solely on the basis of goals scored, the model keeps score in units of home team win probability. In this model the score changes both continuously with time and with jumps corresponding to changes in the goal and manpower differential state of the game.

# 7 Win Probability Added

A common problem faced by managers and analysts of team sports is how to assess the marginal contribution of an individual player to team performance, and hockey is no exception. The NHL's adopted approach to this problem is via the "+/-" statistic that, whenever an even strength, shorthanded or empty net goal is scored, awards "+1" to all players on the ice for the scoring team excluding the goalie, while penalizing the skaters on the ice for the scored-upon team by "-1" (again excluding the goalie). These +/- points are aggregated across games and seasons, and used as information to discern a player's net contribution to winning.

Various authors have attempted to estimate a hockey player's marginal contribution to net goals scored using statistical models. MacDonald (2010) and Schuckers (2011) applied regression methods to shift-level data (that is, data collected while the same collection of players is on ice) to produce "adjusted +/- statistics." Gramacy, Jensen and Taddy (2013) used logistic models to estimate the amount of "credit" individual players should receive when goals are scored, while Thomas et al (2013) employed competing hazards models to develop player ratings. Regardless of approach, previous research has found it difficult to obtain accurate estimates of a player's marginal contribution to net goals for non-elite players.

Given that goals are relatively rare in hockey (recall that both home and away teams average fewer than three goals per game), while the overarching objective of a team is to win, a better measure would score players at all moments they are on the ice during a game in terms of their contribution to the probability that their team wins the game. Summing such incremental win probability contributions over time can be thought of as assessing the statistical number of games each player wins for the team. Such "win probability added" statistics have existed for some time in baseball (Tango, Lichtman and Dolphin, 2007; Winston, 2009), and are routinely updated in real-time at fangraphs.com, for example. In-game win probability models of the sort summarized in Figure 5 provide a basis for meeting this objective in hockey, as we will now show.

Excluding the goalie, let

$$\xi_i^H(t) = \begin{cases} 1 & \text{home team player } i \text{ on ice at time } t \\ \\ 0 & \text{otherwise} \end{cases} \tag{10}$$

and

$$n_H(t) = \sum_i \xi_i^H(t) = \text{ number home team skaters on ice at time } t. \tag{11}$$

Further, define $\omega(t)$ as the home team win probability at elapsed time $t$ in a game as computed by applying the results of our model displayed in Figure 2 to the actual in-game goal and manpower differential state transitions, an example of which is the solid curve in Figure 5. With these definitions, the *win probability added* (*WPA*) for the $i^{th}$ player on the home team equals

$$WPA_i^H = \int_0^\tau \frac{\xi_i^H(t)}{n_H(t)}\, d\omega(t), \tag{12}$$

while the analogous measure for the $j^{th}$ player on the away team equals

$$WPA_j^A = -\int_0^\tau \frac{\xi_j^A(t)}{n_A(t)}\, d\omega(t) \tag{13}$$

where $n_A(t)$ is the number of away team skaters on the ice at time $t$, and $\tau$ now refers to the duration of the hockey game (including overtime if applicable).

We do not attempt to apply *WPA* to shootouts that occur when overtime fails to determine game winner (the +/- statistic also ignores shootouts).

The logic behind equations (12-13) is simple: $d\omega(t)$ is the instantaneous change in the home team win probability at time $t$ in the game (and $-d\omega(t)$ is the instantaneous change in win probability for the away team). Presuming that all skating players are equally responsible for their team's change in win probability (a similar assumption applies to standard +/- statistics, namely, that all skating players share responsibility for goals scored or allowed), the ratio $\xi_i^H(t)/n_H(t)$ ($\xi_j^A(t)/n_A(t)$) captures home player $i$'s (away player $j$'s) share of this change. Cumulating these win probability changes over all time for each player in the game yields equations (12-13). Note that unlike the +/- statistics, win probability added includes skaters on power plays in addition to even strength or short-handed situations, and as the number of skaters on ice changes for either team, the "shared responsibility" for changes in win probability is divided accordingly.

To illustrate, consider again the Dallas Stars at New York Islanders game reported in Table 3 and Figure 5. Table 4 reports the +/- results for this game. Josh Bailey was the only Islander with a positive +/- statistic, while Doug Weight and James Wisniewski both have +/- statistics of $-1$. For the Stars, Brenden Morrow, Mike Ribeiro and Stephane Robidas each have +/- statistics equal to +2. Of note is that Morrow, Wisnewski and Weight were selected as the "three stars" of this particular game (http://www.nhl.com/scores/htmlreports/20102011/GS020015.HTM), which in the case of Wisnewski and Weight is perhaps surprising given their negative +/- statistics.

**New York Islanders**

| Player | G | A | +/- |
|---|---|---|---|
| J. Bailey LW | 0 | 1 | 1 |
| B. Comeau LW | 1 | 0 | 0 |
| M. Eaton D | 0 | 0 | 0 |
| T. Gillies LW | 0 | 0 | 0 |
| T. Hunter RW | 0 | 0 | 0 |
| M. Jurcina D | 0 | 0 | -1 |
| Z. Konopka C | 0 | 0 | 0 |
| A. MacDonald D | 0 | 0 | -1 |
| R. Martinek D | 0 | 1 | 0 |
| M. Mottau D | 0 | 0 | -1 |
| M. Moulson LW | 1 | 0 | -2 |
| N. Niederreiter RW | 0 | 0 | 0 |
| F. Nielsen C | 0 | 1 | 0 |
| P. Parenteau RW | 0 | 2 | -1 |
| J. Sim RW | 0 | 0 | -1 |
| J. Tavares C | 0 | 0 | -1 |
| D. Weight C | 1 | 2 | -1 |
| J. Wisniewski D | 1 | 1 | -1 |

**Dallas Stars**

| Player | G | A | +/- |
|---|---|---|---|
| K. Barch RW | 0 | 0 | -1 |
| Jamie Benn LW | 1 | 0 | 0 |
| A. Burish RW | 0 | 0 | 1 |
| T. Daley D | 0 | 1 | 0 |
| L. Eriksson LW | 0 | 0 | 1 |
| M. Fistric D | 0 | 0 | -1 |
| N. Grossmann D | 0 | 1 | 1 |
| B. Morrow LW | 2 | 0 | 2 |
| J. Neal LW | 0 | 0 | 1 |
| M. Niskanen D | 0 | 0 | 1 |
| S. Ott C | 0 | 1 | 0 |
| T. Petersen C | 0 | 0 | 0 |
| M. Ribeiro C | 0 | 2 | 2 |
| B. Richards C | 0 | 2 | 0 |
| S. Robidas D | 1 | 0 | 2 |
| K. Skrastins D | 0 | 0 | 1 |
| B. Sutherby C | 0 | 0 | 0 |
| T. Wandell C | 0 | 0 | 0 |

Table 4: Plus/minus statistics for Dallas at New York Islanders from http://scores.espn.go.com/nhl/boxscore?gameId=301009012.

Table 5 reports our win probability added measures for each player in this game as calculated via equations (12-13) with $\omega(t)$ following Figure 5. Comparing to the +/- statistics, note that while Josh Bailey of the Islanders has the highest *WPA* for the Islanders (just as he had the highest +/- statistic), Doug Weight and James Wisnewski have the third and fourth highest *WPA* measures for the Islanders. This is a reassuring result, in that these two players were both recognized as stars of this game, and their positive *WPA* measures reinforce that view. Similarly, Brenden Morrow of the Stars, another three star selection, also recorded positive *WPA*. Recalling that Morrow was one of three Dallas skaters who received +/- statistics of +2, it is interesting that the other two (Mike Ribeiro and Stephane Robidas) recorded *negative* *WPA* measures by our calculations. This example illustrates that *WPA* and +/- are not capturing the same thing, and though admittedly anecdotal, it is reassuring that the selection of the three stars of the game is more consistent with *WPA* than with +/-.

| Islanders | WPA | Mins | Stars | WPA | Mins |
|---|---|---|---|---|---|
| Josh Bailey | 0.1097 | 22.15 | Trevor Daley | 0.0700 | 25.78 |
| Trent Hunter | 0.0660 | 17.35 | Adam Burish | 0.0530 | 12.35 |
| Doug Weight | 0.0549 | 20.33 | Steve Ott | 0.0381 | 16.47 |
| James Wisniewski | 0.0353 | 24.17 | Brad Richards | 0.0348 | 24.58 |
| Nino Niederreiter | 0.0305 | 11.45 | Brenden Morrow | 0.0292 | 19.97 |
| Frans Nielsen | 0.0303 | 18.25 | Toby Petersen | 0.0272 | 11.38 |
| Andrew MacDonald | -0.0007 | 23.58 | James Neal | 0.0271 | 20.42 |
| Zenon Konopka | -0.0031 | 11.38 | Matt Niskanen | 0.0229 | 17.75 |
| Trevor Gillies | -0.0049 | 1.95 | Karlis Skrastins | 0.0168 | 19.60 |
| Matt Moulson | -0.0100 | 21.52 | Loui Eriksson | 0.0091 | 23.58 |
| Jon Sim | -0.0113 | 12.52 | Jamie Benn | -0.0097 | 13.23 |
| Milan Jurcina | -0.0119 | 18.18 | Brian Sutherby | -0.0129 | 6.55 |
| Pa Parenteau | -0.0373 | 17.08 | Stephane Robidas | -0.0141 | 26.05 |
| John Tavares | -0.0467 | 5.30 | Tom Wandell | -0.0147 | 10.25 |
| Radek Martinek | -0.0470 | 20.60 | Krystofer Barch | -0.0236 | 6.00 |
| Mark Eaton | -0.0575 | 18.82 | Mike Ribeiro | -0.0256 | 22.20 |
| Blake Comeau | -0.0609 | 20.90 | Nicklas Grossman | -0.0696 | 19.33 |
| Mike Mottau | -0.0884 | 21.10 | Mark Fistric | -0.1051 | 12.13 |

Table 5: Individual win probability added for the skaters in Dallas at NY Islanders.

There is another property of *WPA* evident in Table 5. Note that summing the *WPA* measures for the Islanders yields $-0.05$ while doing the same for Dallas gives $+0.05$. This is not a coincidence. Inspection of Figure 2 shows that at the start of a game (with both goal and manpower differentials of zero), the home team has a 55% probability of winning. At the end of overtime in a tie game, the home team has a 50% probability of winning. Thus, the change in home team win probability from the start of a game to the end of unresolved overtime is given by $0.50 - 0.55 = -0.05$ (and by symmetry, the change in away team win probability equals $+0.05$). If the home team wins a hockey game (whether in regulation or overtime), the win probability jumps to 1, and thus the sum of *WPA* for the home team must equal $1-0.55 = 0.45$; conversely if the home team loses, the sum of individual player *WPA* will equal $-0.55$. Summarizing, at the end of a game, the sum of player $WPA$ for the home team equals

$$WPA^H = \sum_i WPA_i^H = \begin{cases} 0.45 & \text{home team wins} \\ -0.05 & \text{OT tie} \\ -0.55 & \text{home team loses} \end{cases} \tag{14}$$

while for the visiting team, we have

$$WPA^A = \sum_j WPA_j^A = \begin{cases} 0.55 & \text{away team wins} \\ 0.05 & \text{OT Tie} \\ -0.45 & \text{away team loses} \end{cases} . \tag{15}$$

For a given team, over the course of an entire regular season, let $\#HW$, $\#HT$, $\#HL$, $\#AW$, $\#AT$, and $\#AL$ reflect the number of games the team wins at home, ties after overtime at home, loses at home, wins away, ties after overtime away, and loses away respectively. Summing all player *WPA*s over all games yields team *WPA* for the entire season, and making use of equations (14-15) we see that

$$
\begin{aligned}
\text{Team } WPA &= 0.45 \times \#HW - 0.05 \times \#HT - 0.55 \times \#HL \\
&\quad +0.55 \times \#AW + 0.05 \times \#AT - 0.45 \times \#AL \\
&= 0.45 \times \#HW - 0.55 \times (41 - \#HW) + 0.55 \times \#AW - 0.45 \times (41 - \#AW) \\
&= \#HW + \#AW - 41 \\
&= \text{ Wins Over 500.}
\end{aligned} \tag{16}
$$

Noting that there are 41 home and 41 away games for each team in an 82 game NHL regular season, we can write

$$\#HT = 41 - \#HW - \#HL \tag{17}$$

and

$$\#AT = 41 - \#AW - \#AL \tag{18}$$

which enables an alternative expression for team *WPA*:

$$
\begin{aligned}
\text{Team } WPA &= 0.45 \times \#HW - 0.05 \times \#HT - 0.55 \times \#HL \\
&\quad +0.55 \times \#AW + 0.05 \times \#AT - 0.45 \times \#AL \\
&= 0.5 \times \#HW - 0.5 \times \#HL - 0.5 \times 41 \\
&\quad +0.5 \times \#AW - 0.5 \times \#AL + 0.5 \times 41 \\
&= 0.5 \times (\text{Wins } - \text{ Losses}).
\end{aligned} \tag{19}
$$

Equations (16) and (19) complete the connection from individual player *WPA*s for each team in each game to end-of-season winning records. Also, summing individual player *WPA*s across an entire season record each player's contribution to Wins Over 500.

## 8 Conclusions

We have developed a new Markov model for NHL hockey incorporating manpower differential in addition to goal differential into the state space. The model was calibrated based on more than 17.7 million second-by-second state observations during regulation time of every regular-season NHL game played between the fall of 2008 and spring of 2012. The model clarifies the tradeoffs between goals and penalties by showing that scoring a goal is worth more than a four person change in manpower differential. We showed how the model can provide in-game win probability assessments that in turn can be used to evaluate the contributions of individual hockey players to the probability that their team wins the game. We also showed how win probability added cumulates across players and games to equal the number of wins over average for any given team.

Future research opportunities include deploying our approach to estimating win probability added to all players (excluding goalies) in all games across a given season to see how the resulting player evaluation measures compare to those already computed. As with adjusted +/-, it might also prove important to adjust *WPA* to account for linuep correlation and variable opponent skill. Our method currently does not apply to goaltenders, but surely goalie contributions to win probability matter and should be accounted for in some way. There are also some timing issues that perhaps deserve additional attention. For example, consider a penalty that is successfully killed by two different shifts. In the present approach, only the shift on ice when the penalty expires receives the jump in win probability that accompanies the favorable transition in manpower differential, even though the first unit surely deserves some credit for killing the penalty. Nonetheless, we believe that our model does suggest new ways to think about the relative gains and losses associated with goals and penalties while offering a direct method for valuing the incremental contributions of individual players to the probability their team wins hockey games.

**References**

Beaudoin, David and Tim B. Swartz (2010). Strategies for pulling the goalie in hockey. *The American Statistician* 64(3): 197-204.

Berry, Scott M. (2000) My triple crown - First leg: pulling the goalie. *Chance*, 13(3): 56-57.

Bishop, Yvonne M. M., Stephen E. Fienberg, and Paul W. Holland (1975). *Discrete Multivariate Analysis: Theory and Practice*. Cambridge, MA: MIT Press.

Burke, Brian (2009). NHL In-Game Win Probability. http://www.advancednflstats.com/2009/04/nhl-in-game-win-probability.html (accessed February 21, 2014).

Buttrey, Samuel E., Alan R. Washburn, and Wilson L. Price (2011). Estimating NHL Scoring Rates. *J. Quantitative Analysis in Sports* 7: Iss. 3, Art. 24.

Berger, Jonah, and Devin Pope (2011). Can Losing Lead to Winning? *Management Science* 57(5): 817-827.

Erkut, Erhan (1987) Note: More on Morrison and Wheat's "Pulling the Goalie Revisited." *Interfaces* 17(5): 121-123.

Macdonald, Brian (2011). A Regression-Based Adjusted Plus-Minus Statistic for NHL Players. *J. Quantitative Analysis in Sports* 7: Iss. 3, Art. 4.

Morrison, Donald G. (1976). On the optimal time to pull the goalie: A Poisson model applied to a common strategy used in ice hockey." *TIMS Studies in Management Science* 4: 137-144.

Moskowitz, Toby J. and L. Jon Werthem (2011). *Scorecasting: The Hidden Influences Behind How Sports Are Played And Games Are Won.* New York: Crown Archtype.

Mullet, Gary M. (1977). Simeon Poisson and the National Hockey League. *The American Statistician* 31(1): 8-12.

Nydick, Robert L. and Howard J. Weiss (1989). More on Erkut's "More on Morrison and Wheat's 'Pulling the Goalie Revisited'." *Interfaces* 19(5): 45-48.

Tango, Tom M., Mitchel G. Lichtman and Andrew E. Dolphin (2007). *The Book: Playing The Percentages In Baseball.* Washington, DC: Potomac Books, Inc.

Thomas, A. C., Samuel L. Ventura, Shane Jensen, and Stephen Ma (2013). Competing process hazard function models for player ratings in ice hockey. *Annals of Applied Statistics* 7(3): 1249-1835

Schuckers, Michael E., Dennis F. Lock, Chris Wells, C. J. Knickerbocker, and Robin H. Lock (2011). National Hockey League skater ratings based upon all on-ice events: an adjusted minus/plus probability (AMPP) approach." Working Paper: http://myslu.stlawu.edu/~msch/sports/ LockSchuckersProbPlusMinus113010.pdf (accessed February 21, 2014).

Stern, Hal (1994). A Brownian motion model for the progress of sports scores. *Journal of the American Statistical Association* 89(427): 1128-1134.

Washburn, Alan (1991). Still more on pulling the goalie. *Interfaces* 21(2): 59-64.

Winston, Wayne L. (2009). *Mathletics: How Gamblers, Managers, and Sports Enthusiasts Use Mathematics in Baseball, Basketball and Football.* Princeton: Princeton University Press.

Zaman, Z. (2001). Coach Markov Pulls Goalie Poisson. *Chance* 14(2): 31-35.



# 9 Statistical Appendix

## 9.1 State Occupancy: Modeled and Observed

The Markov framework used to model win probability can also be employed to model the probability that a game is in any particular goal/manpower differential state at any point in time. To see this, let $p(x, y, t)$ denote the probability that after $t$ seconds of play (that is, at *elapsed* time $t$ as opposed to $t$ seconds from the end of regulation time), the home team leads by $x$ goals while enjoying a manpower advantage of $y$. At the start of the game ($t = 0$), the score is tied while the teams are at even strength, thus $p(0, 0, 0) = 1$. From this time on, the game states evolve in accord with the same state-dependent transition rates $\lambda_{xy}^{x'y'}$ discussed earlier, which leads to the state evolution equation

$$\frac{dp(x,y,t)}{dt} = \sum_{(x',y')\neq(x,y)} \lambda^{xy}_{x'y'} p(x',y',t) - \left(\sum_{(x',y')\neq(x,y)} \lambda^{x'y'}_{xy}\right) p(x,y,t). \quad (20)$$

$$x = -4, ..., 4;\ y = -2, ..., 2;\ 0 < t \leq 3600$$

in conjunction with the aforementioned initial condition $p(0,0,0) = 1$ (which of course implies $p(x,y,0) = 0$ for all $(x,y) \neq (0,0)$), as well as the boundary conditions $p(\pm 5, y, t) = 0$ to remain consistent with our five goal mercy rule. As an example, Figure A1 reports the modeled (via equation (20)) and observed fraction of time over all 17.7M seconds of hockey in our data that hockey games were in played at even strength ($y = 0$) with home minus away goal differentials ranging from $x = -2$ to $x = 2$. The correspondence between observed and modeled state occupancy probability is remarkable, again considering that the observed state occupancy fractions were not in any way taken into account when estimating the model parameters (the state transition rates). This is thus a pure test of the Markov theory, and with the exception of the final ten minutes of tie games at even strength (some of which could be explained by teams pulling the goalie and successfully causing a tie), the visual fit is remarkable. Beyond goodness of fit, equation (20) will prove important in assessing the model-implied variability of observed win probabilities in various game states, as will be discussed below.

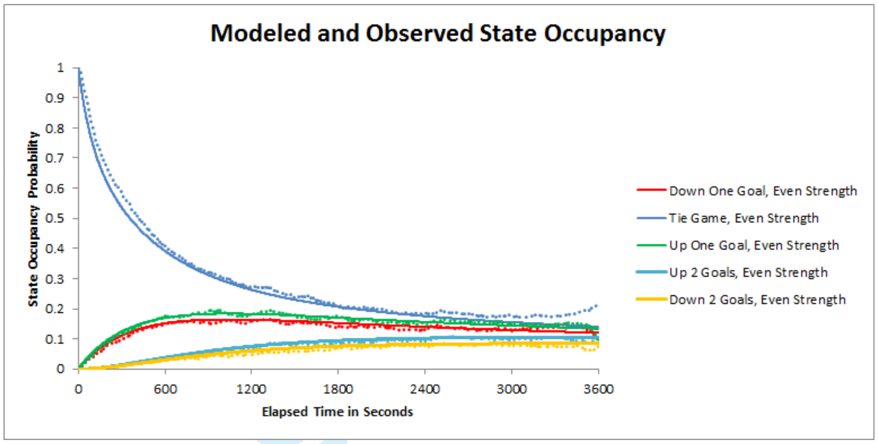


Figure A1: Comparing observed and modeled state occupancy probability by goal differential for games at even strength.

## 9.2 Win Probabilities: Modeled and Observed

Given our proposed use of the in-game win probability model expressed via equations (3)-(8) for use with actual game transitions to compute win probability added for individual players, it is important that the modeled and observed win probabilities are close enough to enable trust in the model. But how close is close enough? The answer to this depends upon the variability in *observed* win probabilities one expects *based on the model.*

Before addressing this issue formally, consider again Figure A1 which reports even strength state occupancy probabilities for goal differentials ranging from $-2$ to $+2$. At all times over the first 5 minutes of the game, the probability of being in an even strength tie game exceeds 50%. Thus the state occupancy probabilities for all other goal/manpower differential states must sum to less than 50%! Now consider how one computes empirically observed state-dependent win probabilities from our data. Focusing on the five minute mark over all 4,920 games in our data set, one asks: how many of these games were in the goal/manpower state of interest, and of those games observed in the relevant state, in how many did the home team win?

Focusing on the 5 minute mark, we would expect about half of all 4,920 games to be tie games at even strength. By contrast, based on equation (20), one would only expect about 1.4% of games to be at even strength but with the home team leading by two goals, and one would only expect about 1.1% of all games to be at even strength with the home team trailing by two goals. One should therefore expect that observed win probabilities 5 minutes into the game would exhibit much less variability given a tie game at even strength than observed win probabilities for games at even strength where the home team is up or down by two goals, as there are so many more tie games than 2 goal leads at the five minute mark. And indeed, this is exactly what is observed. Figure A2 shows the observed number of wins over the observed number of games for even strength games with goal differentials $x$ ranging from $-2$ to $+2$. Note how much tighter the observed win probabilities hug the Markov model for tie games.

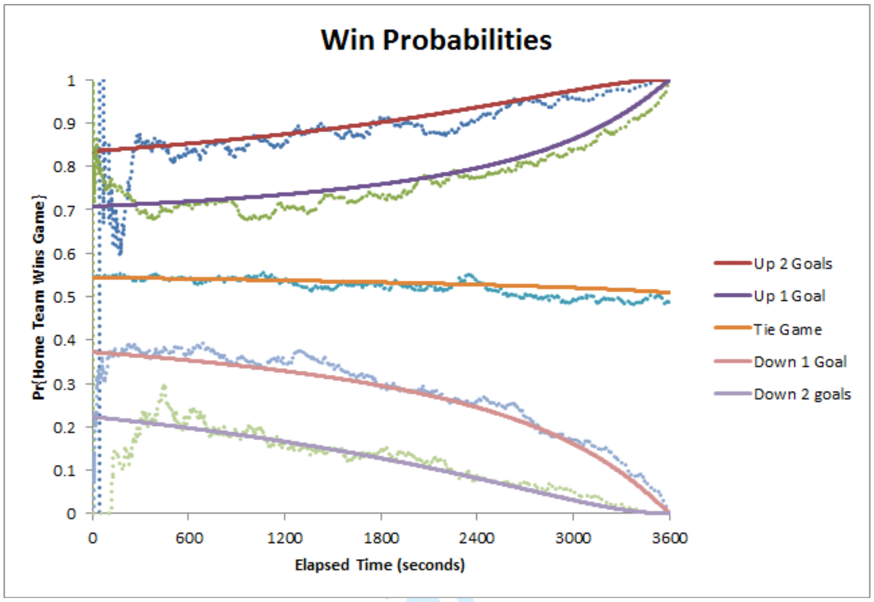


Figure A2: Comparing observed and modeled home win probability by goal differential for games at even strength.

The situation in uneven manpower differential states is more complicated. First, such states are very rare relative to game play at even strength. This is consistent with our model given the transition rates estimated from the data. Second, because uneven manpower situations are relatively rare, the observed fraction of the time games are in such situations is highly variable relative to their expected values (but very low in actual value). Both of these points are illustrated in Figure A3. The solid lines show the state occupancy probabilities resulting from equation (20) when the home team is up or down one goal while playing up or down a man. Of these four states, none are expected to occur more than 2.5% of the time at any point during regulation, while over most of the game, the chance of occupying one of these four states is between 1% and 2%. The dotted lines show the observed fraction of the 4,920 games in our data that, in any second, fall into one of these four states. While the observed data generally follow expectations, their fluctuations relative to the modeled values are much wilder than what is seen for the even strength state occupancies reported in Figure A1.

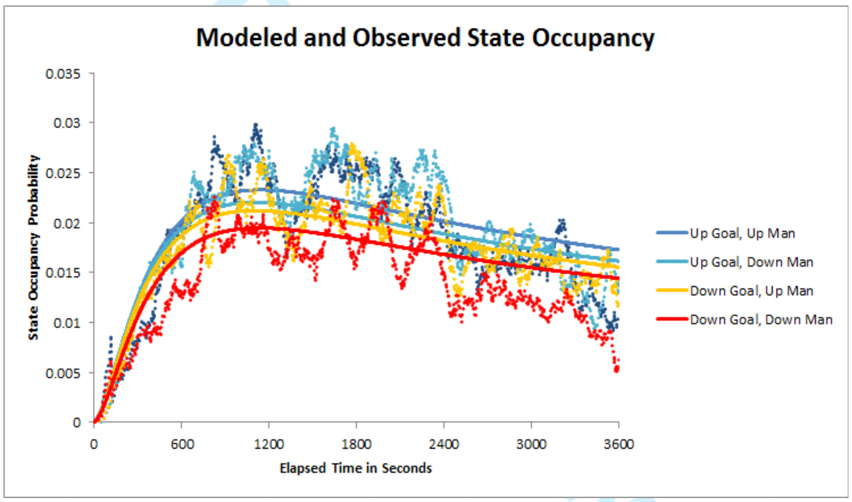


Figure A3: Comparing observed and modeled state occupancy probability by goal differential for games with manpower differentials of ±1.

Now consider modeled versus observed win probabilities for states with different goal/manpower differentials. The observed win probabilities, again while generally following the model, exhibit great variability. But, such fluctuations should be expected based on the model, for computing the empirical win probability requires dividing the observed number of times a the home team wins in a particular state by the number of games in which the game was in that state. With such low likelihoods of even being in uneven manpower states (as is clear from Figure A3), the empirical win probabilities are ratios of very small numbers; indeed, the state occupancy probabilities dictate the denominators in such expressions. Thus, the model predicts that the variability in empirical win probabilities will be large relative to the modeled win probability in states with low occupancy probabilities. Figure A4 displays exactly such behavior for states with goal/manpower differentials from $-1$ to $+1$. Compared to Figure A2 that showed modeled and observed win probabilities for the much more prevalent even strength states, the fluctuations in observed win probabilities relative to the model are much more violent.

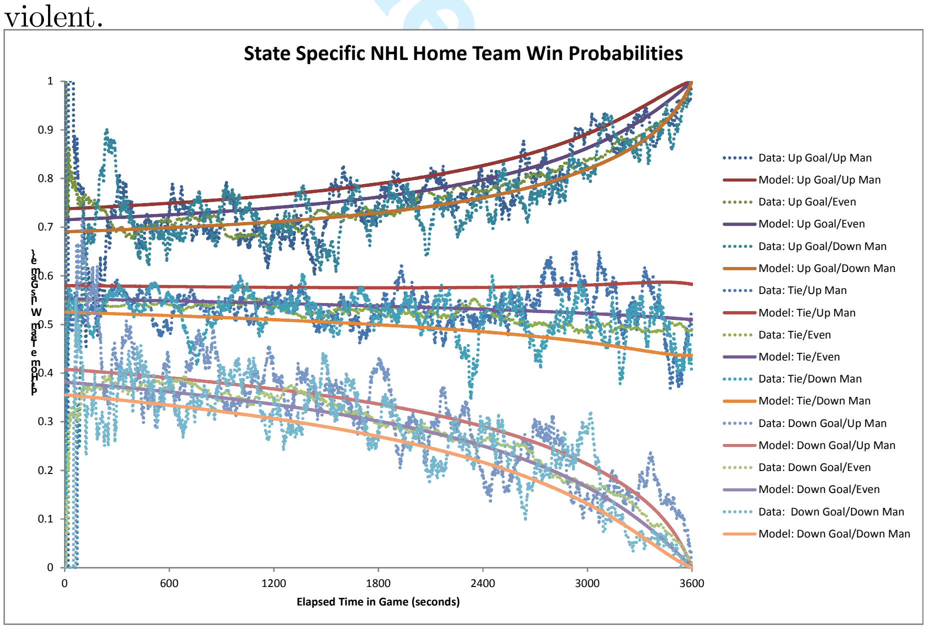


Figure A4: Comparing observed and modeled home win probability by goal differential for games with manpower differentials of $-1, 0, 1$.

## 9.3 Model-Implied Variability in Observed Win Probabilities

Having illustrated informally how model-derived state occupancy probabilities influence the variability in observed win probabilities, we will produce a formula for the variance of observed win probabilities conditional on the underlying model. With such an estimate in hand, it is then possible to see whether observed win probabilities are sufficiently close to the model (e.g. within a 95% confidence interval). We will focus on point-in-time variability, meaning that the procedure to be described can be applied to compare observed and modeled win probabilities at any specified time during a game, and will thus suppress the time dimension $t$ in our notation. Consider a dataset of $n$ games with goal/manpower differential state information recorded at all times for each game (in our case, $n = 4,920$). For some fixed point-in-time during all of these $n$ games, and for any game state $s = (x, y)$, define the random variable $G_s$ as the number of the $n$ games that are in state $s$, and define the random variable $V_s$ as the number of these $G_s$ games in which the home team wins (that is, the number of $v$ictories). We take $G_s$ as binomially distributed with $n$ trials and success probability

$$p_s = p(x, y, \cdot) \tag{21}$$

where the right hand side is computed from equation (20) at whatever time is of interest for the $(x, y)$ state in question. Consequently, we have

$$E(G_s) = np_s \tag{22}$$

and

$$Var(G_s) = np_s(1 - p_s). \tag{23}$$

We model $V_s$ as conditionally binomial (given $G_s$) with $G_s$ trials and success probability

$$v_s = w(x, y, \cdot) \tag{24}$$

where the right hand side is computed from equation (3), again at whatever time is of interest for the $(x, y)$ state in question. Recognizing that the home team can only win a game in state $s$ if the game is actually in state $s$ (which occurs with probability $p_s$) and the home team wins (which occurs with probability $v_s$), the mean and variance of $V_s$ follow

$$E(V_s) = np_sv_s \tag{25}$$

and

$$Var(V_s) = np_sv_s(1 - p_sv_s). \tag{26}$$

Also, note that

$$\begin{aligned} E(G_s V_s) &= E_{G_s}[E(G_s V_s | G_s)] \\ &= E_{G_s}[G_s^2 v_s] \\ &= v_s E(G_s^2) \end{aligned} \tag{27}$$

while

$$\begin{aligned} E(G_s)E(V_s) &= np_s \times np_s v_s \\ &= v_s E(G_s)^2 \end{aligned} \tag{28}$$

and thus the covariance $Cov(G_s, V_s)$ is given by

$$\begin{aligned} Cov(G_s, V_s) &= E(G_s V_s) - E(G_s)E(V_s) \\ &= v_s E(G_s^2) - v_s E(G_s)^2 \\ &= v_s Var(G_s) > 0. \end{aligned} \tag{29}$$

The empirical win probability estimator for games in state $s$ is then given by

$$W_s = \frac{V_s}{G_s}. \tag{30}$$

Our goal is to arrive at the mean and especially the variance of $W_s$ for purposes of comparing our win probability model ($w(x, y, t)$) to empirical observations. We ignore the possibility that $G_s = 0$ which will pose no problems providing we focus on game states that occur with sufficiently high probability (below we show that $p_s > 0.01$ suffices), and use the delta approximation method (Bishop, Fienberg, Holland, 1975)) to arrive at very simple results in closed form. We then relax the assumptions imposed by the delta method, and use simulation to generate numerical results. Foreshadowing our results, we will demonstrate that as long as the game state of interest occurs at least 1% of the time, the simulated variance of $W_s$ is essentially indistinguishable from that implied by the delta method.

Informally, the delta method approximation presumes that given an estimator $\widehat{\boldsymbol{\theta}}$ that follows a multivariate normal distribution with mean $\theta$ and covariance matrix $\Sigma$, a new random variable $f(\widehat{\boldsymbol{\theta}})$ will, providing $f$ is differentiable, be approximately multivariate normal with mean $f(\theta)$ and variance $f'(\theta)\,\Sigma\, f'(\theta)^T$ ($f'(\theta)$ is the Jacobian matrix of first derivatives). In our application, we take $\widehat{\boldsymbol{\theta}} = (V_s, G_s)$, $\theta = (E(V_s), E(G_s))$, and

$$\Sigma = \begin{pmatrix} Var(V_s) & v_s Var(G_s) \\ v_s Var(G_s) & Var(G_s) \end{pmatrix}. \tag{31}$$

We expect this to work well under the usual conditions for approximating binomial random variables by their normal counterparts. The function $f(\widehat{\boldsymbol{\theta}})$

is given by $W_s = V_s/G_s$, and thus

$$f'(\theta) = \left(\frac{dW_s}{dV_s}, \frac{dW_s}{dG_s}\right) = \left(\frac{1}{G_s}, -\frac{V_s}{G_s^2}\right). \tag{32}$$

Applying the delta method, we approximate the mean of the empirical win probability estimator by

$$E(W_s) = \frac{E(V_s)}{E(G_s)} = \frac{np_s v_s}{np_s} = v_s, \tag{33}$$

which shows that under the delta method, this estimator is unbiased. The approximate variance of the empirical win probability estimator is (the reader can verify the result)

$$Var(W_s) = \left(\frac{1}{E(G_s)}, -\frac{E(V_s)}{E(G_s)^2}\right) \begin{pmatrix} Var(V_s) & v_s Var(G_s) \\ v_s Var(G_s) & Var(G_s) \end{pmatrix} \begin{pmatrix} \frac{1}{E(G_s)} \\ -\frac{E(V_s)}{E(G_s)^2} \end{pmatrix} = \frac{v_s(1-v_s)}{np_s}. \tag{34}$$

This is an intuitively satisfying result; were the number of games in state $s$ fixed in advance at, say, $n_s$, then the usual estimate for the variance of a binomial proportion, $v_s(1-v_s)/n_s$, would apply. The result above replaces the denominator with the *expected* number of games in state $s$, with the expectation deriving from the state occupancy model of equation (20). Equations (33) and (34) provide us with the benchmark needed to compare observed and expected win probabilities, for one can simply ask whether the observed fraction of games in a given state in which the home team wins falls within +/- two standard deviations of the empirical win probability estimator, that is, whether the observed result falls within +/- twice the square root of equation (34) from $v_s$.

Equation (34) highlights the important role of the state occupancy probability in determining the model-implied variability in observed win probabilities. In infrequently occupied states (very low values of $p_s$), the variance of the empirical win probability estimator will be inflated, while in high probability states, the reverse occurs. The model win probability enters as $v_s(1 - v_s)$ as is familiar from standard sampling theory, so 50/50 games will have the highest sampling variability in observed win probabilities. What is important is that our model provides both win probabilities and state occupancy probabilities, and both must be taken into account.

The results above rely on the assumptions of the delta method, so before continuing it is worth checking to see if these assumptions are terribly restrictive. We do this easily via simulation: for any given $(p_s, v_s)$ pair, we first simulate a binomial random variable with $4,920$ trials (this is the number of games in our data) and success probability $p_s$; the result represents the number of games observed (in the simulation) in state $s$. Call this $g_s$. We then simulate a second binomial random variable with $g_s$ trials and success probability $v_s$; the result represents the number of games in state $s$ that are won by the home team. Call this $h_s$ (for home victories in state $s$, with apologies for running out of symbols). Finally, we take the ratio $h_s/g_s$ as the empirical win probability in the simulation. We replicate this 2,500 times, which provides a simulation-based estimate of the sampling distribution for the empirical win probability in our data from which we can compute its mean and variance.

Figure A5 presents the standard deviation of the empirical win probability that results from this procedure, assuming game state occupancies ranging from 1% to 25% and associated win probabilities between 0 and 1 along with results of using the formulas based upon the delta method. As is clear from the figure, the simple approximation provides excellent results compared to the simulation over this range.

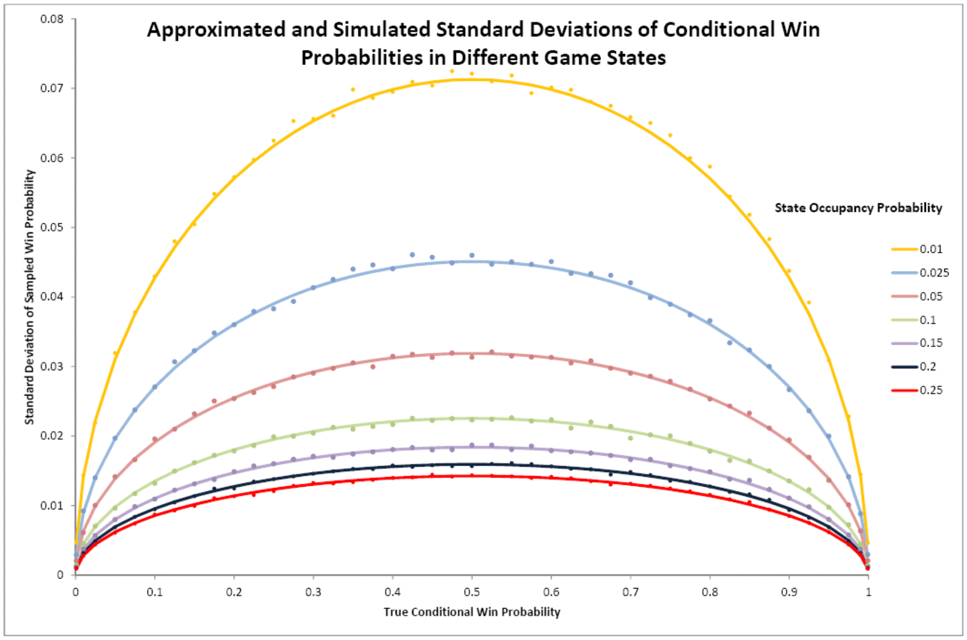


Figure A5: Approximated (solid lines) and simulated (dotted lines) standard deviations of conditional win probabilties.

The fit of our model to the data can now be illustrated graphically. In Figures A6-A8 we plot the observed and modeled win probabilities for three different cases: even strength in a tie game (this is the highest probability state); home team leads by a goal while playing with a man advantage; and home team down a goal and a man. We also show point-in-time-wise 95% confidence intervals based on +/- two standard deviations as computed from the square root of equation (34). For tie games at even strength, the model fit is truly excellent. The fit is not as good in the other game states, but generally speaking the actual observed win probabilities are not far from where they should be. For states with very low occupancy probabilities, the fluctuations in observed win probabilities seem violent, but in light of their model-implied variability, such fluctuations should be expected.

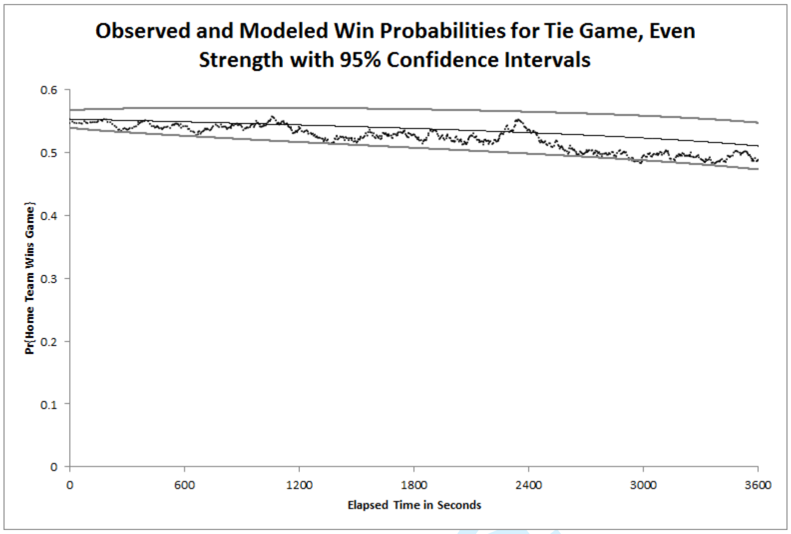


Figure A6: Observed and modeled win probabilities with approximate 95% confidence intervals for tie games at even strength.

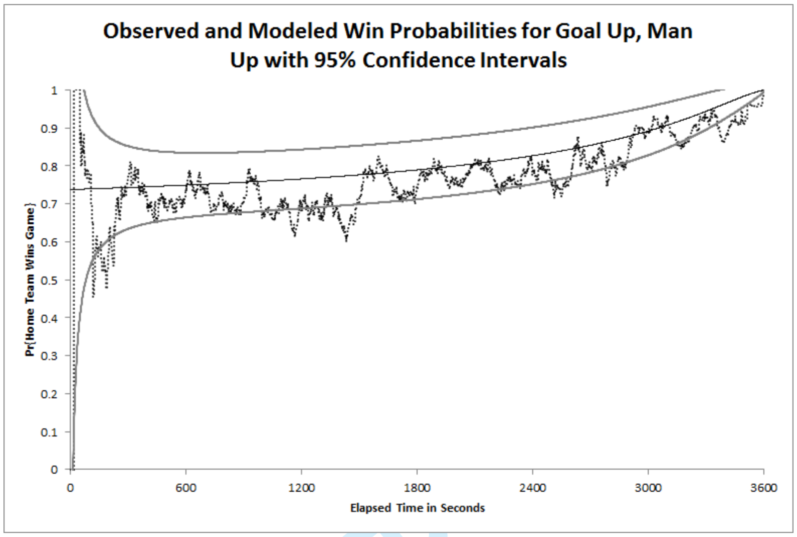


Figure A7: Observed and modeled win probabilities with approximate 95% confidence intervals for games with goal and manpower differentials of 1.

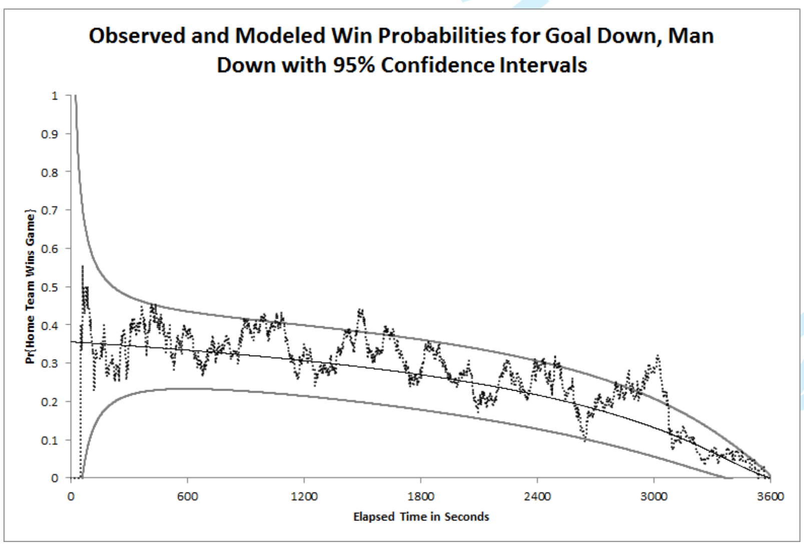


Figure A8: Observed and modeled win probabilities with approximate 95% confidence intervals for games with goal and manpower differentials of $-1$.