Assessing the Offensive Productivity of NHL Players Using In-game Win Probabilities

Stephen Pettigrew

Harvard University

9TH ANNUAL MIT SLOAN SPORTS ANALYTICS CONFERENCE
FEBRUARY 27-28, 2015 BOSTON CONVENTION AND EXHIBITION CENTER
presented by ESPN

Imagine that you're an NHL GM.

You want to add some goal scoring talent to your roster.

Two potential acquisitions

24 years old
15 goals in 2014-2015
0.26 goals per game (career)
48.1% Corsi-tied (career)

24 years old
16 goals in 2014-2015
0.30 goals per game (career)
47.6% Corsi-tied (career)

9th ANNUAL MIT SLOAN SPORTS ANALYTICS CONFERENCE
FEBRUARY 27-28, 2015 BOSTON CONVENTION AND EXHIBITION CENTER
presented by ESPN

ticketmaster

42 ANALYTICS

Which player should you trade for?

The one who scores in high pressure situations.

or

The one who is more clutch

Outline of this talk

An NHL Win Probability Model

Narrative Building using Win Probabilities

Measuring Offensive Productivity with Added Goal Value

Outline of this talk

An NHL Win Probability Model

Narrative Building using Win Probabilities

Measuring Offensive Productivity with Added Goal Value

An NHL Win Probability Metric

- Hockey journalists and statisticians currently lack many empirical tools available in other sports

  One such tool is a metric for calculating second-by-second win probabilities

- The metric introduced here:
  - is the only such metric currently available
  - incorporates powerplay information in a novel way
  - provides a flexible framework which can unify the work being done on puck possession, zone starts, spatial data, etc.

Intuition about the metric

$$
\begin{aligned}
P_t(w) \quad = \quad & P_t(w|\delta_t + 1) \cdot \Lambda(\gamma_h \cdot \nu_t) + \\
& P_t(w|\delta_t - 1) \cdot \Lambda(\gamma_a \cdot \nu_t) + \\
& P_t(w|\delta_t)(1 - \Lambda(\gamma_h \cdot \nu_t))(1 - \Lambda(\gamma_a \cdot \nu_t))
\end{aligned}
$$

When the teams are playing at full-strength, the vector ν_t is all zeros. This means that the metric simplifies to:

$$P_t(w) = P_t(w|\delta_t)$$

Intuition about the metric

$$
\begin{aligned}
P_t(w) &= P_t(w|\delta_t + 1) \cdot \Lambda(\gamma_h \cdot \nu_t) + \\
&\quad P_t(w|\delta_t - 1) \cdot \Lambda(\gamma_a \cdot \nu_t) + \\
&\quad P_t(w|\delta_t)(1 - \Lambda(\gamma_h \cdot \nu_t))(1 - \Lambda(\gamma_a \cdot \nu_t))
\end{aligned}
$$

If the home team is on a powerplay then:

Intuition about the metric

$$\begin{aligned} P_t(w) &= P_t(w|\delta_t + 1) \cdot \Lambda(\gamma_h \cdot \nu_t) + \\ &\quad P_t(w|\delta_t - 1) \cdot \Lambda(\gamma_a \cdot \nu_t) + \\ &\quad P_t(w|\delta_t)(1 - \Lambda(\gamma_h \cdot \nu_t))(1 - \Lambda(\gamma_a \cdot \nu_t)) \end{aligned}$$

If the home team is on a powerplay then:

- $\Lambda(\gamma_h \cdot \nu_t)$ is the probability that the home team scores a PP goal

Intuition about the metric

$$\begin{aligned} P_t(w) &= P_t(w|\delta_t + 1) \cdot \Lambda(\gamma_h \cdot \nu_t) + \\ & \quad P_t(w|\delta_t - 1) \cdot \Lambda(\gamma_a \cdot \nu_t) + \\ & \quad P_t(w|\delta_t)(1 - \Lambda(\gamma_h \cdot \nu_t))(1 - \Lambda(\gamma_a \cdot \nu_t)) \end{aligned}$$

If the home team is on a powerplay then:

- $\Lambda(\gamma_h \cdot \nu_t)$ is the probability that the home team scores a PP goal
- $\Lambda(\gamma_a \cdot \nu_t)$ is the probability that the away team scores a SH goal

Intuition about the metric

$$
\begin{aligned}
P_t(w) \quad = \quad & P_t(w|\delta_t + 1) \cdot \Lambda(\gamma_h \cdot \nu_t) + \\
& P_t(w|\delta_t - 1) \cdot \Lambda(\gamma_a \cdot \nu_t) + \\
& P_t(w|\delta_t)(1 - \Lambda(\gamma_h \cdot \nu_t))(1 - \Lambda(\gamma_a \cdot \nu_t))
\end{aligned}
$$

If the home team is on a powerplay then:

- $\Lambda(\gamma_h \cdot \nu_t)$ is the probability that the home team scores a PP goal
- $\Lambda(\gamma_a \cdot \nu_t)$ is the probability that the away team scores a SH goal
- $(1 - \Lambda(\gamma_h \cdot \nu_t))(1 - \Lambda(\gamma_a \cdot \nu_t))$ is the probability that neither team scores before they return to even strength

A flexible framework

$$
\begin{aligned}
P_t(w) \quad = \quad & P_t(w|\delta_t+1)\cdot \Lambda(\gamma_h\cdot\nu_t)+ \\
& P_t(w|\delta_t-1)\cdot \Lambda(\gamma_a\cdot\nu_t)+ \\
& P_t(w|\delta_t)(1-\Lambda(\gamma_h\cdot\nu_t))(1-\Lambda(\gamma_a\cdot\nu_t))
\end{aligned}
$$

Everything in red represents the probability that a goal is scored by one team or the other in some period of time.

The framework here allows for any statistics to be included, as long as you have a model that predicts how the stat impacts goal-scoring rates

- Which players are on the ice
- Zone starts
- Corsi and Fenwick stats
- Spatial data

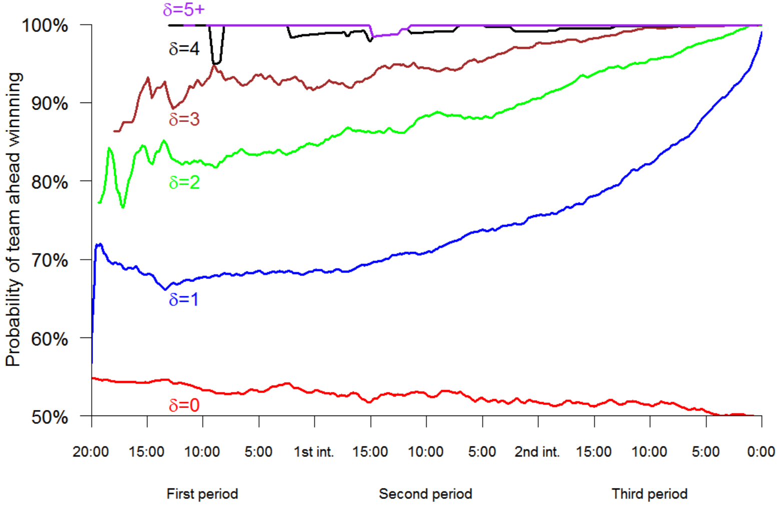
Estimating $P_t(w|\delta_t)$

I used empirical data to estimate $P_t(w|\delta_t)$

- Data from all regular season games from 2005/2006 through 2012/2013
- Over 9000 games

I then used a weak Bayesian prior to smooth out the data

Estimating $P_t(w|\delta_t)$

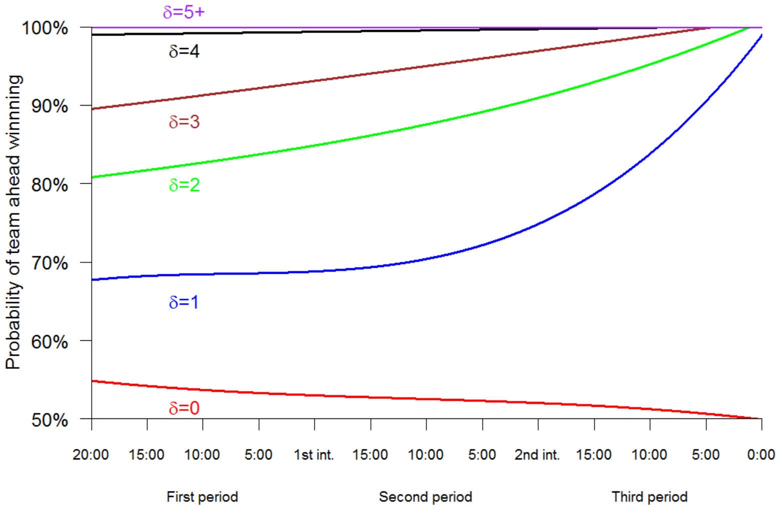


MIT SLOAN SPORTS ANALYTICS CONFERENCE
FEBRUARY 27-28, 2015 BOSTON CONVENTION AND EXHIBITION CENTER
presented by ESPN

ticketmaster®

42 ANALYTICS

Estimating $P_t(w|\delta_t)$



Outline of this talk

An NHL Win Probability Model

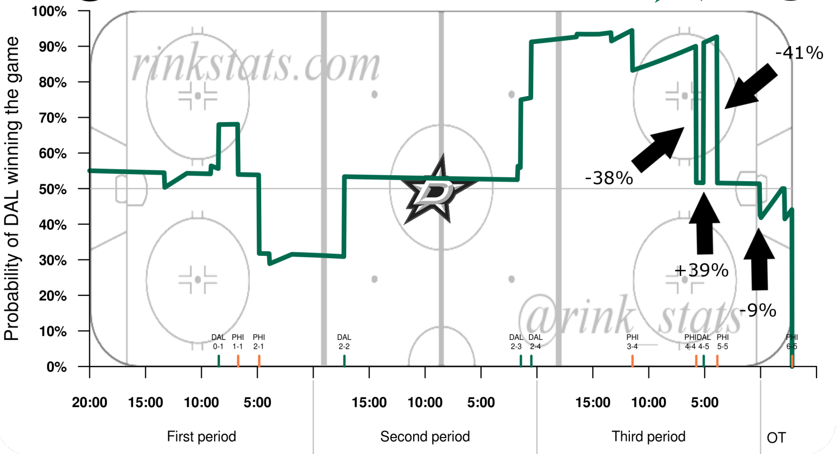
Narrative Building using Win Probabilities

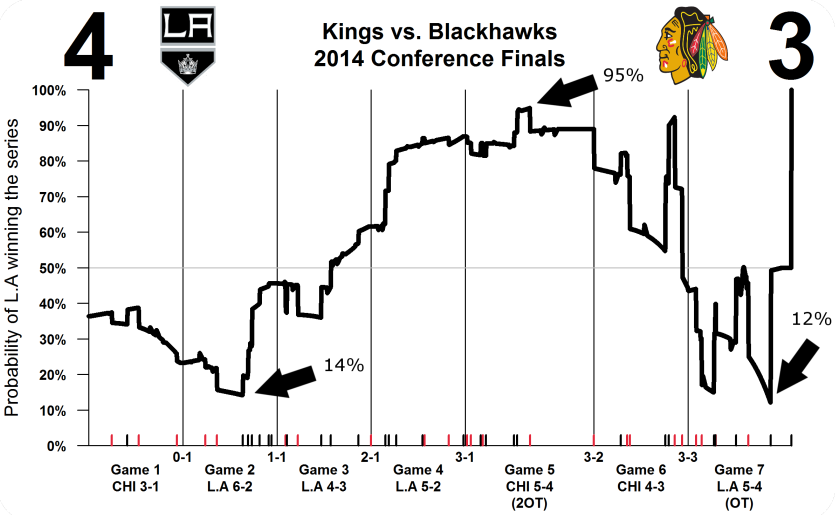
Measuring Offensive Productivity with Added Goal Value

6

Flyers at Stars
Saturday, October 18, 2014

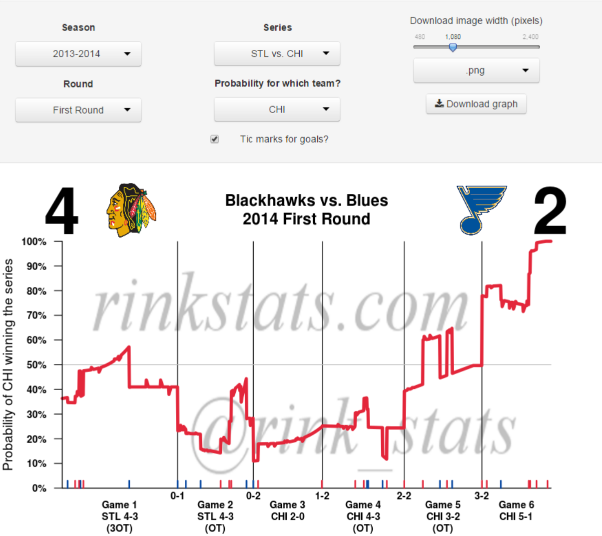
5





4
3
Kings vs. Blackhawks
2014 Conference Finals
Probability of L.A winning the series
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
95%
14%
12%
0-1
1-1
2-1
3-1
3-2
3-3
Game 1
CHI 3-1
Game 2
L.A 6-2
Game 3
L.A 4-3
Game 4
L.A 5-2
Game 5
CHI 5-4
(2OT)
Game 6
CHI 4-3
Game 7
L.A 5-4
(OT)

Win probability apps



See the apps at:
winprobs.rinkstats.com

Also featured on
war-on-ice.com

9th ANNUAL MIT SLOAN SPORTS ANALYTICS CONFERENCE
FEBRUARY 27-28, 2015 BOSTON CONVENTION AND EXHIBITION CENTER
presented by ESPN

ticketmaster®

42 ANALYTICS

Outline of this talk

An NHL Win Probability Model

Narrative Building using Win Probabilities

Measuring Offensive Productivity with Added Goal Value

Added Goal Value

Not all goals are created equal.

A 6^{th} goal in a 5-0 game is less valuable to a team than an overtime winner.

Traditional statistics like total goals masks this fact and treats every goal as equally important.

Added Goal Value (AGV) uses win probabilities to measure the significance of each goal. Players with high AGV can be thought of as being more "clutch."

The math behind Added Goal Value

$$AGV_{Crosby} = \sum_{k=1}^{K} \Delta_k - \frac{\sum_{j=1}^{J} \Delta_j}{J} \cdot K$$

$$\begin{aligned} \Delta_k &= P_{t_k+1}(w) - P_{t_k}(w) \\ \Delta_j &= P_{t_j+1}(w) - P_{t_j}(w) \end{aligned}$$

$\Delta_k = P_{t_k+1}(w) - P_{t_k}(w)$: change in win probability from a goal scored by Sidney Crosby

The math behind Added Goal Value

$$AGV_{Crosby} = \sum_{k=1}^{K} \Delta_k - \frac{\sum_{j=1}^{J} \Delta_j}{J} \cdot K$$

$$\begin{aligned} \Delta_k &= P_{t_k+1}(w) - P_{t_k}(w) \\ \Delta_j &= P_{t_j+1}(w) - P_{t_j}(w) \end{aligned}$$

$\Delta_k = P_{t_k+1}(w) - P_{t_k}(w)$: change in win probability from a goal scored by Sidney Crosby

$\Delta_j = P_{t_j+1}(w) - P_{t_j}(w)$: change in win probability from a goal scored another player

9th ANNUAL MIT SLOAN SPORTS ANALYTICS CONFERENCE FEBRUARY 27-28, 2015 BOSTON CONVENTION AND EXHIBITION CENTER presented by ESPN

ticketmaster®

42 ANALYTICS

Understanding Added Goal Value

$$AGV_{Crosby} = \sum_{k=1}^{K} \Delta_k - \frac{\sum_{j=1}^{J} \Delta_j}{J} \cdot K$$

Understanding Added Goal Value

$$AGV_{Crosby} = \sum_{k=1}^{K} \Delta_k - \frac{\sum_{j=1}^{J} \Delta_j}{J} \cdot K$$

$\sum_{k=1}^{K} \Delta_k$: total added win percentage for all goals scored by Crosby

Understanding Added Goal Value

$$AGV_{Crosby} = \sum_{k=1}^{K} \Delta_k - \frac{\sum_{j=1}^{J} \Delta_j}{J} \cdot K$$

$\sum_{k=1}^{K} \Delta_k$: total added win percentage for all goals scored by Crosby

$\frac{\sum_{j=1}^{J} \Delta_j}{J}$: average increase in win probability for all goals (about 17%)

Understanding Added Goal Value

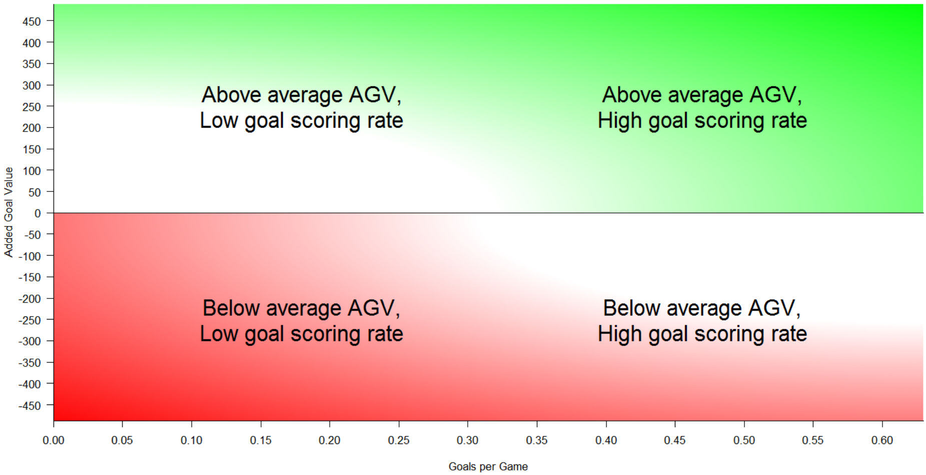
$$AGV_{Crosby} = \sum_{k=1}^{K} \Delta_k - \frac{\sum_{j=1}^{J} \Delta_j}{J} \cdot K$$

$\sum_{k=1}^{K} \Delta_k$: total added win percentage for all goals scored by Crosby

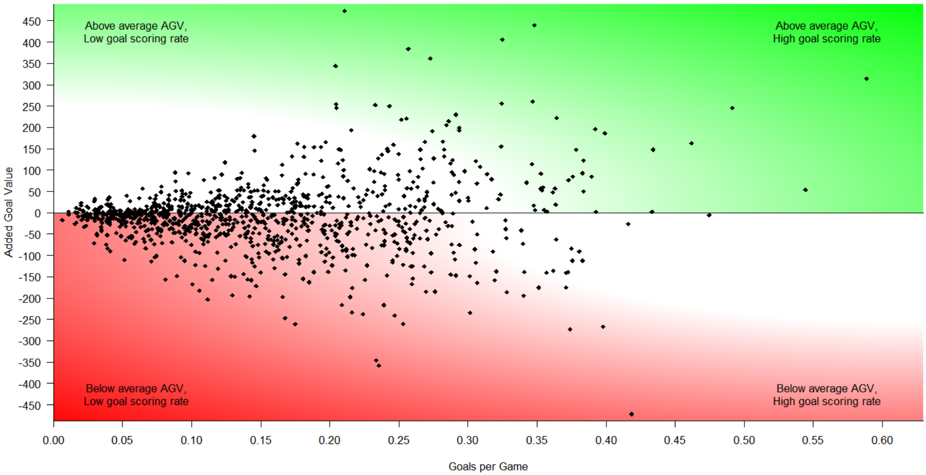
$\frac{\sum_{j=1}^{J} \Delta_j}{J}$: average increase in win probability for all goals (about 17%)

$\frac{\sum_{j=1}^{J} \Delta_j}{J} \cdot K$: expected increase in win probability resulting from the K goals scored by Crosby

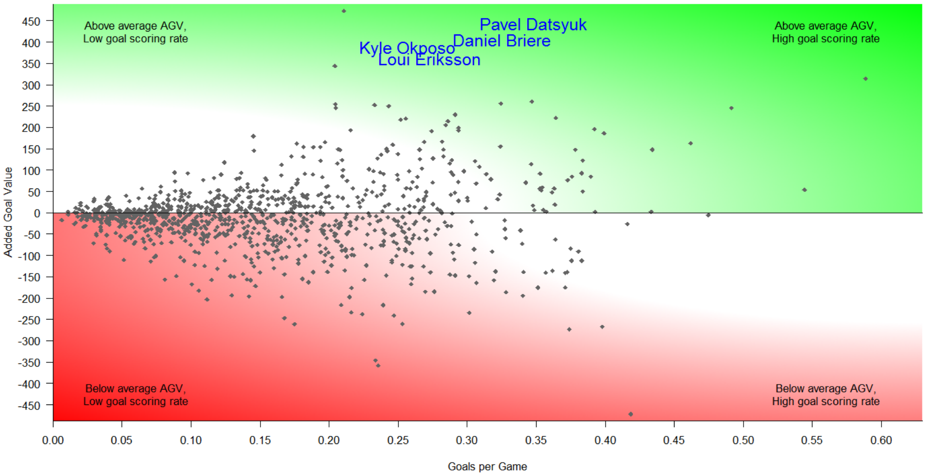
AGV since 2005/2006 season



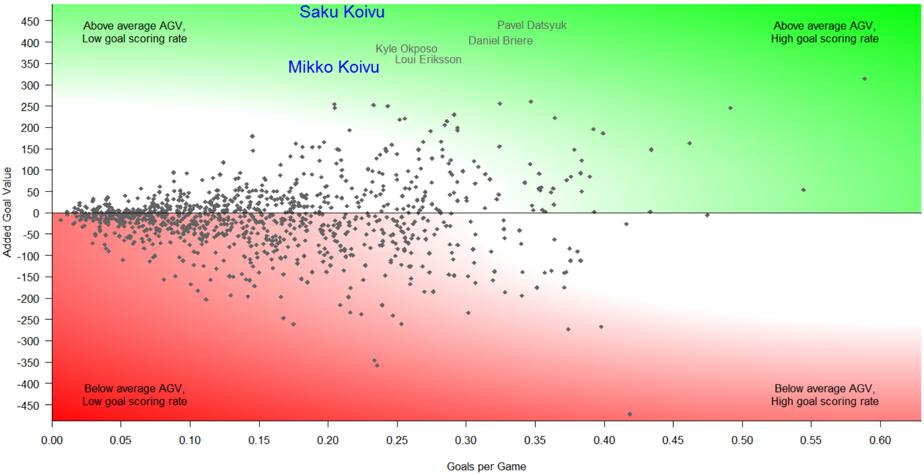
AGV since 2005/2006 season



AGV since 2005/2006 season



AGV since 2005/2006 season

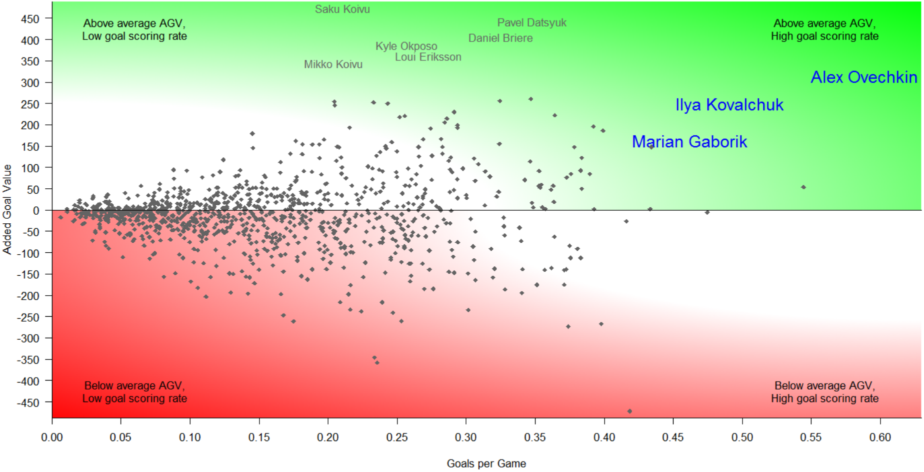


9TH ANNUAL MIT SLOAN SPORTS ANALYTICS CONFERENCE
FEBRUARY 27-28, 2015 BOSTON CONVENTION AND EXHIBITION CENTER
presented by ESPN

ticketmaster®

42 ANALYTICS

AGV since 2005/2006 season

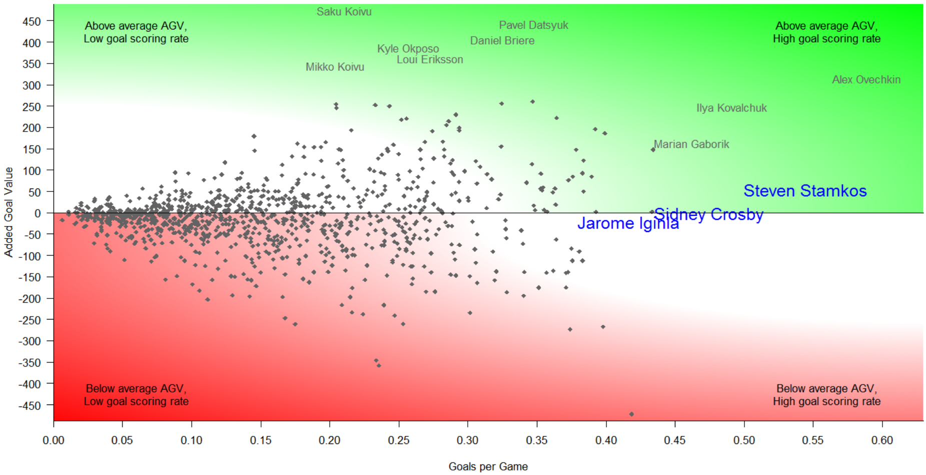


9TH ANNUAL MIT SLOAN SPORTS ANALYTICS CONFERENCE
FEBRUARY 27-28, 2015 BOSTON CONVENTION AND EXHIBITION CENTER
presented by ESPN

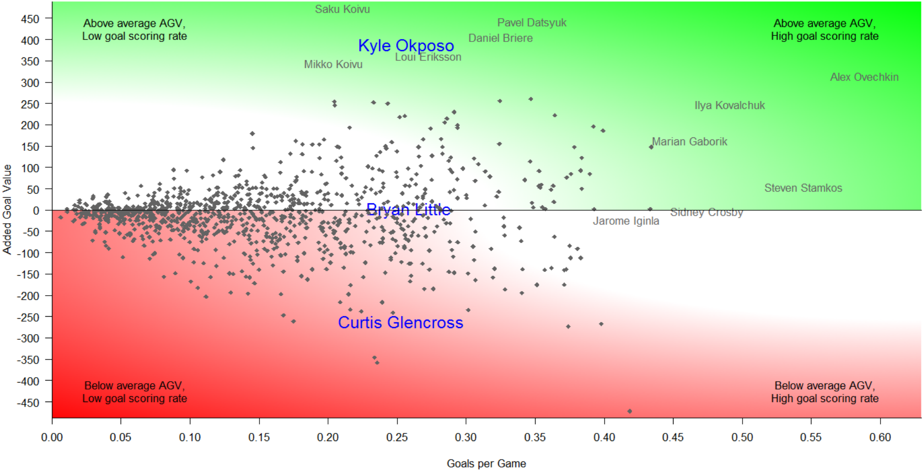
ticketmaster®

42ANALYTICS

AGV since 2005/2006 season



AGV since 2005/2006 season

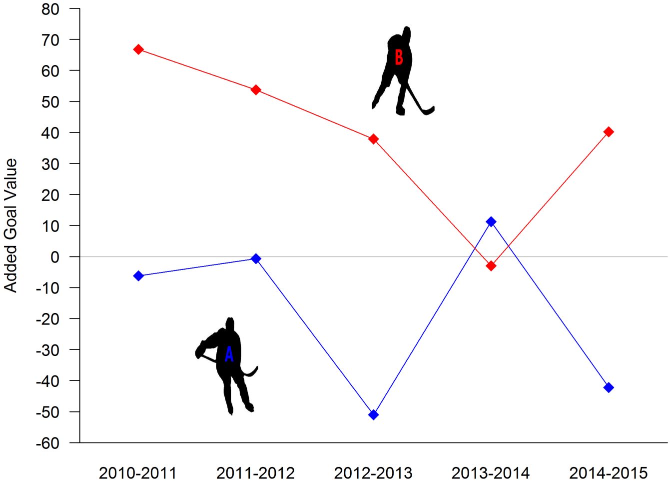


Two potential acquisitions

24 years old
15 goals in 2014-2015
0.26 goals per game (career)
48.1% Corsi-tied (career)

24 years old
16 goals in 2014-2015
0.30 goals per game (career)
47.6% Corsi-tied (career)

AGV for Players A and B



Who to acquire?

-42.2 AGV (2014-2015)
-88.9 AGV (career)

+40.2 AGV (2014-2015)
+198.9 AGV (career)

Next steps

- Enrich the win probability metric by adding more statistics
- Stanley Cup winning probabilities; playoff qualification probabilities
- Adapt AGV to shootout
- Add assists to AGV metric
- AGV app

Thanks for listening!

rinkstats.com
stephenpettigrew.com
@rink_stats