

Chapter Three Notes

1. Modeling Standing Points in the NHL

From the book: p.184 (footnote 23) *The outcome in hockey – or standing points – was regressed on goals scored and goals allowed. This simple regression reveals that each goal scored is worth 0.31 standing points. A goal allowed is worth -0.31 standing points. This model was estimated with team data that began with the 1983-83 season and ended in the 2007-08 campaign. Complete details on this regression can be found in Berri and Brook (2009) and at stumblingonwins.com.*

The model is published in Berri, David J. and Stacey Brook (2010). “On the Evaluation of the “Most Important” Position in Professional Sports.” forthcoming **Journal of Sports Economics**.

The article reports the following model of standing points in the NHL.

$$\text{Standing Points} = a_0 + a_1 \text{GS} + a_2 \text{GA} + e_{it}$$

Where GS = Goals Scored
GA = Goals Against

**OLS Regression Results:
Standings Points 1983/84 – 2007/08**

Variable	Coefficient	t-Statistic
Constant*	80.367	23.283
Goals For*	0.314	46.728
Goals Against*	-0.308	-46.992
D08*	9.548	6.026
D07*	9.780	6.253
D06*	9.723	5.990
D04*	5.403	3.100
D03*	5.701	3.557
D02*	4.582	2.703
D01*	4.558	2.742
D00*	4.568	2.905
D99	0.543	0.316
D98	0.541	0.328
D97	0.426	0.286
D96	0.332	0.223
D95	0.397	0.215
D94	0.291	0.164
D93	0.132	0.091
D92	0.193	0.137
D91	0.203	0.127
D90	0.108	0.076
D89	0.084	0.060
D88	0.096	0.065
D87	0.114	0.075
D86	-0.010	-0.006
D85	0.024	0.015
<hr/>		
R-squared =	0.926	
Adjusted R-squared =	0.923	
F-statistic =	293.801	
Prob(F-statistic) =	0.000	

n = 609, White Heteroskedasticity-Consistent Standard Errors & Covariance

* statistically significant at 1% level.

Additional Notes on Model

- This model was estimated with team data that began with the 1983-83 season and ended in the 2007-08 campaign. In all we had 609 team observations. Year dummies were employed, and there were only significant after the 1999-00 season. White Heteroskedasticity-Consistent Standard Errors & Covariance were employed and the model explained 92% of the variation in standing points. According to Hockey-Reference.com – the source of our hockey data -- this is the first season save percentage is recorded.
- This model is used in our evaluation of goalies (see book for details).

2. Modeling the Salary of Goalies in the NHL

From the book: p. 185 (footnote 31) *A study of unrestricted free agent goalies is challenging. There simply are not many goalies each year that sign a new free agent contract. Plus, the data on these free agents (as Stacey Brook discovered, the co-author of the study our discussion is based upon) is not particularly easy to find. Given these limitations, the sample consisted of 33 unrestricted free agents at the goalie position from 2004 to 2008. To qualify for the sample a goalie had to have played at least two seasons prior to signing the contract (if one only focused on a single lag the sample would have increased to forty). Salary data came from USA Today.com (content.usatoday.com/sports/hockey/nhl/salaries/default.aspx). For details on the specific model employed one is referred to Berri and Brook (2009) and stumblingtonwins.com.*

The model is published in Berri, David J. and Stacey Brook (2010). "On the Evaluation of the "Most Important" Position in Professional Sports." forthcoming **Journal of Sports Economics**.

$$\text{Log}(\text{Salary}_{ij}) = \gamma_0 + \gamma_1\text{PROD} + \gamma_3\text{LAGPROD} + \gamma_4\text{AGE} + \gamma_5\text{SQAGE} + \gamma_6\text{MP} + \gamma_7\text{LAGMP} + \gamma_8\text{POP} + \gamma_9\text{D03} + \gamma_{10}\text{D04} + \gamma_{11}\text{D06} + e_{ij}$$

Where

Prod = is productivity from the season before signing a free agents contract. Productivity can be defined in terms of save percentage, goals against average (GAA), or wins

LagProd = Prod from the two seasons before signing a free agent contract

MP = Minutes played

LAGMP = Minutes played from two seasons ago

POP = Market size for team that signs free agent goalie

D03, D04, and D06 are year dummies

Additional details reported in table and notes below.

Dependent Variable: Log of Salary
 Model estimated with three different definitions of productivity
 Ordinary Least Squares (White Heteroskedasticity-Consistent Standard Errors & Covariance)
 (t-statistics reported beneath each coefficient)

Independent Variables			
Save Percentage, last season	17.320**		
	2.219		
Save Percentage, two years ago	27.577**		
	2.365		
GAA, last season		-0.189	
		-0.657	
GAA, two years ago		-0.188	
		-0.475	
Wins, last season			0.009
			0.344
Wins, two years ago			-0.023
			-0.781
Age	0.836	0.791	0.780
	1.455	1.343	1.480
Age, squared	-0.013***	-0.012	-0.011
	-1.539	-1.371	-1.487
Minutes played, last season	0.00042*	0.00053*	0.00049
	2.901	2.982	1.471
Minutes played, two seasons ago	0.00022**	0.00017	0.00036
	2.159	1.426	1.336
Population	-9.501E-09	-4.984E-09	-9.478E-09
	-0.518	-0.206	-0.388
D2006	-0.242	-0.240	-0.307
	-0.640	-0.539	-0.738
D2004	-0.512	-0.579	-0.610
	-1.327	-1.289	-1.352
D2003	-0.308	-0.563	-0.568
	-0.830	-1.257	-1.426
Constant	-41.413**	-0.581	-0.486
	-2.386	-0.061	-0.054
Adjusted R-squared	0.48	0.29	0.29
Observations	33	33	33

* statistically significant 1%; ** statistically significant 5%; *** statistically significant at the 10% level

We also estimate the model with current productivity. The results are reported below:

Dependent Variable: Log of Salary
 Estimating salary model with three different definitions of current performance of goalie
 Ordinary Least Squares (White Heteroskedasticity-Consistent Standard Errors & Covariance)
 (t-statistics reported beneath each coefficient)

Independent variables	Eq. 4d	Eq. 4e	Eq. 4f
Save Percentage, current season	8.018		
	0.734		
GAA, current season		-0.265	
		-0.842	
Wins, current season			-0.005
			-0.195
Age	0.578	0.660	0.587
	1.396	1.610	1.542
Age, squared	-0.009	-0.010	-0.009
	-1.376	-1.606	-1.540
Minutes played	0.00049*	0.00049*	0.00056**
	3.737	4.013	2.095
Market size	1.02E-08	1.15E-08	8.12E-09
	0.507	0.598	0.384
Constant term	-3.866	2.691	3.246
	-0.385	0.391	0.509
Adjusted R-squared	0.27	0.28	0.27
Observations	33	33	33

* statistically significant 1%; ** statistically significant 5%; *** statistically significant at the 10% level

Additional Notes on Salary Model

- Here are some details on the model we used to examine salaries paid to goalies. Previously we noted studies of salaries in baseball and basketball. In each of these sports there isn't an issue of sample size. A study of free agent goalies, though, is much more difficult. There simply are not many goalies each year that sign a new free agent contract. Plus, the data on these free agents is not particularly easy to find. Given these limitations, our sample consisted of 33 unrestricted free agents at the goalie position from 2004 to 2008. To qualify for our sample you had to have played at last two seasons prior to signing the contract (if we only focused on a single lag our sample would have increased to forty).
- Our list of independent variable begins with performance data (save percentage, wins, goals against), which was taken from Hockey-Reference.com. We also considered the age of the goalie and age squared. Age squared was included because we expect that a goalie will initially improve as he gets older, and then at some point, age reduces the ability of a goalie to play. So the coefficient on age squared should be negative. The estimated coefficients indicated this relationship, although the results were not generally statistically significant. As we did in our study of salaries in the NBA, the proxy for injury is minutes played. We also considered games played. The results were comparable. For market size we considered the size of the metropolitan area where the

team is located. For cities in the United States we used data from the U.S. Census Bureau (2000 census). Data for Canadian cities was taken from Statistics Canada [www.statcan.gc.ca] and this was from their 2006 census. We also considered a separate dummy variable for cities located in Canada and this was insignificant.

- Regressing current salary on current performance fails to find a link between what a goalie is doing today and what he is being paid today. We also considered a much larger data set. Salary data was collected from USA Today on goalies that played at least 2500 minutes in a single season from 2000-01 to 2007-08. We then looked at the relationship between the pay a goalie received and his current save percentage. Across 181 goalie observations we found no statistical relationship between current pay and current save percentage. This larger data set did not consider when the goalie signed his contract or his free agent status when he signed the contract.