

Chapter Two Notes

1. Modeling the Salaries of NBA Free Agents

From the book: p. 178 (footnote 22) *The salary model is an updated version of what was reported in Berri, Brook, and Schmidt (2007). The model specifically examined 337 free agents who signed multiyear contracts from 2001 to 2008. Salary data was taken from both USA Today.com and the Web site of Patricia Bender. Details of this model can be found at www.stumblingonwins.com.*

The salary model was updated from the model reported in the following paper:
Berri, David J., Stacey L. Brook, and Martin B. Schmidt. (2007) "Does One Simply Need to Score to Score?" **International Journal of Sport Finance**, 2, n4; (November): 190-205.

Here is the specific model estimated.

$$\text{LOG(AVGSAL)} = \gamma_0 + \gamma_1 \text{PTS} + \gamma_2 \text{PPS} + \gamma_3 \text{FT} + \gamma_4 \text{TREB} + \gamma_5 \text{BLK} + \gamma_6 \text{AST} + \gamma_7 \text{STL} + \gamma_8 \text{PF} + \gamma_9 \text{TOV} + \gamma_{10} \text{POP} + \gamma_{11} \text{AGE} + \gamma_{12} \text{D5} + \gamma_{13} \text{D4} + \gamma_{14} \text{D2} + \gamma_{15} \text{D1} + \gamma_{16} \text{GM} + \gamma_{17} \text{DRACE} + \gamma_{18} \text{DSAME} + \gamma_{19} \text{TMWINS} + \gamma_{20} \text{GS/GP} + \varepsilon$$

Where	<u>Player Performance Statistics (last two years, adjusted for position played)</u>
	PTS = Points Scored per-minute
	PPS = Points-per-shot [(PTS-FTM)/FGA]
	FT = Free Throw percentage
	TREB = Rebounds per-minute
	BLK = Blocked shots per-minute
	AST = Assists per-minute
	STL = Steals per-minute
	PF = Personal fouls per-minute
	TOV = Turnovers per-minute
	<u>Non-performance variables</u>
	POP = Market size of city where signing team is located
	AGE = Age of player
	D5 = Dummy variable, center
	D4 = Dummy variable, power forward
	D2 = Dummy variable, shooting guard
	D1 = Dummy variable, point guard
	GM = Games played, last two seasons
	DRACE = Dummy variable, race of player (equal to one if black)
	DSAME = Dummy variable, equal to one if player signed with same team
	TMWINS = Regular season victories by team that employed player previous season
	GS/GP = Ratio of games started to games played
	Year dummy variables were also included

Dependent Variables: Log of Real Average Salary

Ordinary Least Squares Estimation

White Heteroskedasticity-Consistent Standard Errors & Covariance

Sample: 337 free agents signing multi-year contracts from 2001-2008

Variable	Coefficient	Standard Error	t-Statistic	p-value
Constant Term	12.586*	0.628	20.047	0.000
Points Scored	2.195*	0.425	5.159	0.000
Points-per-shot	0.651***	0.369	1.764	0.079
Free Throw Percentage	0.518	0.391	1.326	0.186
Rebounds	3.945*	0.743	5.312	0.000
Blocked Shots	3.450**	1.723	2.002	0.046
Assists	2.615**	1.146	2.282	0.023
Steals	3.046	2.691	1.132	0.258
Personal Fouls	-4.429*	1.417	-3.127	0.002
Turnovers	2.496	3.005	0.831	0.407
Population	7.8E-09	0.000	1.507	0.133
Age of Player	-0.040*	0.009	-4.552	0.000
Dummy variable, center	0.064	0.077	0.833	0.406
Dummy variable, power forward	0.067	0.083	0.812	0.418
Dummy variable, shooting guard	-0.279**	0.108	-2.578	0.010
Dummy variable, point guard	-0.019	0.079	-0.239	0.812
Game Played, last two seasons	0.006*	0.001	4.057	0.000
Dummy variable, race of player	-0.011	0.072	-0.151	0.880
Dummy variable, signed with same team	0.147**	0.057	2.574	0.011
Team wins	0.008*	0.003	2.887	0.004
Ratio of Games Started to Games Played	0.493*	0.094	5.224	0.000
R-squared	0.674			
Adjusted R-squared	0.646			

* - denotes significance at the 1% level

** - denotes significance at the 5% level

*** - denotes significance at the 10% level

Additional notes on the model

- We only considered players who just signed a contract [a practice that follows from Jenkins (1996)]. Lewin and Rosenbaum (2007) recently illustrated why the Jenkins approach is necessary. These authors examined a data set that included all NBA players. The results reported by these authors indicated that scoring totals were the primary determinants of player salary. The results also indicated, though, that shooting efficiency and steals had a negative – and statistically significant -- impact on player salaries. Such a result suggests that players who miss more shots get paid more. Before anyone believes such analysis, though, it's important to note that the data set included players who signed contracts years before the performance data was generated. Furthermore, it appears players were evaluated who were still playing under their rookie contract. The failure to restrict the salary data might explain such odd findings.
- In addition to just signing a contract, we also only included players with significant playing time on the court in the year prior to signing; where “significant playing time” is defined as a player who averaged at least 12 minutes per game and who played in at least 20 games.
- The specific dependent variable is the log of average salary; an average salary that was adjusted for the size of the salary cap in each season.
- Our independent variables included – as detailed in the text -- the per-minute box score statistics tabulated for the player in the two years prior to becoming a free agent. These statistics included both points-per-shot (our measure of shooting efficiency from the field). As noted, points-per-shot is simply double a player's adjusted field goal percentage. The player statistics were also weighted for position played. This process involves calculating the per-minute value of each statistic for the player. We then determined the average per-minute value at each position, as well as the per-minute value across all NBA players. With values in hand, we then subtracted the per-minute position average from each player's per-minute value, and then added back in the average value across all NBA players. For example, Eddy Curry grabbed 0.100 rebounds per-minute in 2004-05. To determine his position adjusted value, we subtract 0.260 – or the number of rebounds an average center grabs per-minute – from Curry's value. We then add back in 0.174, or the average across all players. This gives us a value of 0.014, a value that is now comparable to players at other positions.
- Typically one argues that a coefficient that is significant at the 5% level is statistically significant. Except for shooting efficiency, the player statistics reported as statistically significant in Table 2.3 met the 5% threshold. Shooting efficiency, though, was only significant at the 10% level; and if one only consider what a player did last year – as was done in *The Wages of Wins* -- shooting efficiency becomes insignificant.
- The model considered both turnovers per minute and turnover percentage. Turnover percentage – as detailed at basketball-reference.com – is calculated by dividing turnovers by field goal attempts + 0.44*free throw attempts + turnovers. This numbers is then multiplied by 100. Turnover percentage is essentially an estimate of turnovers per possession. The advantage of using this measure is that it is not highly correlated with points scored per game. The inclusion of turnover percentage, though, still indicated that turnovers and free agent salaries are not statistically related.

- We also used the average value across the two seasons prior to the player signing his new contract. We actually tried a variety of weighting schemes to see if decision-makers valued data from the most recent season more (or less) than the data from two years ago. The results suggested that a simple average was as good a weighting scheme as any scheme we could employ.
- Beyond the player statistics we also considered dummy variables for position played. To avoid perfect multicollinearity, the dummy variable for small forward was omitted. We also included dummy variables for each year (excluding 2001), the race of the player (1 for black), and whether the player signed with the same team. This was included because the Larry Bird exemption allows a team to exceed the salary cap to re-sign one of its own players.
- In addition to dummy variables we also included the number of games the player played across the past two seasons (to control for injury), population in the market where the player was signed, age of the player, the number of games a player's team won, and the ratio of games started to games played. These latter two factors were from the year before the player became a free agent.

2. Modeling the Voting for the All-Rookie Team

p. 180 (footnote 38) *The All-Rookie voting model is an updated version of what was reported in Berri, Brook, and Schmidt (2007). The model examined rookie data from 1995 to 2009. Details of this model can be found at www.stumblingtonwins.com.*

The rookie voting model was updated from the model reported in the following paper:
Berri, David J., Stacey L. Brook, and Martin B. Schmidt. (2007) "Does One Simply Need to Score to Score?" **International Journal of Sport Finance**, 2, n4; (November): 190-205.

Here is the specific model estimated.

$$\text{Log(VP)} = \beta_0 + \beta_1 \text{PTS} + \beta_2 \text{PPS} + \beta_3 \text{FT} + \beta_4 \text{TREB} + \beta_5 \text{BLK} + \beta_6 \text{AST} + \beta_7 \text{STL} + \beta_8 \text{PF} + \beta_9 \text{TOV} + \beta_{10} \text{POP} + \beta_{11} \text{AGE} + \beta_{12} \text{D5} + \beta_{13} \text{D4} + \beta_{14} \text{D2} + \beta_{15} \text{D1} + \beta_{16} \text{GM} + \beta_{17} \text{GS/GP} + \beta_{18} \text{TMWINS} + \beta_{19} \text{DFT} + \varepsilon$$

Where

VP = Voting Points received

Player Performance Statistics (adjusted for position played)

PTS = Points Scored per-minute

PPS = Points-per-shot [(PTS-FTM)/FGA]

FT = Free Throw percentage

TREB = Rebounds per-minute

BLK = Blocked shots per-minute

AST = Assists per-minute

STL = Steals per-minute

PF = Personal fouls per-minute

TOV = Turnovers per-minute

Non-performance variables

POP = Market size of city where team is located

AGE = Age of player

D5 = Dummy variable, center

D4 = Dummy variable, power forward

D2 = Dummy variable, shooting guard

D1 = Dummy variable, point guard

GM = Games played

GS/GP = Ratio of games started to games played

TMWINS = Regular season victories by team that employed player previous season

DFT = Where player was taken in the draft

Dependent Variables: Log of Voting Points
Estimation Method: TOBIT
Sample: 417 rookies from 1994-95 to 2008-09
(t-stats in italics)

Variable	Slope Coefficient	Standard Error	z-statistic	p-value
Points Scored	9.205*	1.001	9.190	0.000
Points-per-shot	-0.132	0.771	-0.170	0.864
Free Throw Percentage	1.001	0.825	1.210	0.225
Rebounds	3.956**	1.691	2.340	0.019
Blocked Shots	1.443	4.142	0.350	0.728
Assists	5.345**	2.423	2.210	0.027
Steals	9.014***	5.357	1.680	0.092
Personal Fouls	-12.905*	2.544	-5.070	0.000
Turnovers	-4.528	4.802	-0.940	0.346
Population	2.1E-08	1.3E-08	1.640	0.101
Age of Player	0.024	0.036	0.670	0.506
Dummy variable, center	-0.332***	0.199	-1.670	0.096
Dummy variable, power forward	0.263	0.186	1.420	0.157
Dummy variable, shooting guard	0.114	0.188	0.610	0.544
Dummy variable, point guard	0.151	0.185	0.820	0.414
Games Played	0.049*	0.005	9.350	0.000
Ratio of Games Started to Games Played	2.031*	0.200	10.150	0.000
Team wins	0.033*	0.005	6.320	0.000
Draft Position	-0.015*	0.004	-3.880	0.000
Observations				
Left censored	115			
Right censored	24			
Uncensored	278			
Total	417			
R-squared	0.724			
Adjusted R-squared	0.710			

* - denotes significance at the 1% level
**_- denotes significance at the 5% level
***_- denotes significance at the 10% level

Additional notes on the model

- Each NBA head coach is asked to vote for the All-Rookie team. Each rookie who receives a first team vote is given two points. A second team selection is worth one point. A coach cannot vote for players on his team, so the maximum voting points a player could receive in the 30-team NBA is 58. From 1995–96 to 2003–04 only 29 teams played in the NBA, so the maximum points was only 56. Prior to 1995–96 only 27 teams played, so maximum points was only 52. The voting data we used were taken from the web site of Patricia Bender.
- The number of maximum voting points a player could receive depends upon the number of teams in the league. Given that the number of teams increased from 27 to 30 in our sample, we adjusted the number of voting points a player received so we could make a

comparison across the years we considered. Specifically, we adjusted upward, so that a player who received 52 voting points in 1995, the maximum possible in a league of 27 teams, was given 58 voting points or the maximum possible in 2005.

- We also included in our study all players who received votes, as well as all rookies who did not receive votes yet who played at least twelve minutes per game and appeared in 41 contests. In other words, we considered rookies with zero votes who conceivably could have been considered for these teams. There is an interesting feature common to both ends of our sample. If a player received the maximum number of votes possible, improvements in performance would not lead to more votes. For those who received zero votes, further declines in performance would not have reduced the number of votes received. Consequently, we estimated a TOBIT or censored model, which one can learn about in any number of econometric textbooks. In all we had 417 observations, of which 277 were uncensored. We would also note that we estimated a semi-logged model (with voting points logged), rather than a linear model. Again, details on the difference would be in any good econometrics book.
- In 1998–99 only 50 games were played. To allow us to compare rookies from this season to the other years we consider, the variables tabulated for players from the 1998-99 season were adjusted. Specifically, we multiplied a factor like games played or team wins by 82/50.
- Like our study of free agents salaries, the statistics were measured on a per-minute basis and adjusted for position played. In *The Wages of Wins* we considered how much of each statistic a player accumulated in his rookie season. In Berri, Brook, and Schmidt (2007) we considered per game averages. The difficult with each approach is that totals and per game statistics are somewhat correlated, and hence the problem of multicollinearity can be a concern. By considering per-minute statistics this problem is mitigated. One should note that the different approaches do yield different results with respect to the statistical significance of everything except points scored and rebounds. The story that points scored dominates the evaluation of coaches, though, remains the same.
- The model considered both turnovers per minute and turnover percentage. Turnover percentage – as detailed at basketball-reference.com – is calculated by dividing turnovers by field goal attempts + 0.44*free throw attempts + turnovers. This numbers is then multiplied by 100. Turnover percentage is essentially an estimate of turnovers per possession. The advantage of using this measure is that it is not highly correlated with points scored per game. The inclusion of turnover percentage, though, still indicated that turnovers and voting for the All-Rookie team were not statistically related.