

PART III



BASKETBALL

This page intentionally left blank

BASKETBALL STATISTICS 101

The Four-Factor Model

For each player and team NBA box scores track the following information:

- two-point field goals made and missed
- three-point field goals made and missed
- free throws made and missed
- personal fouls committed
- assists
- offensive and defensive rebounds
- blocked shots
- turnovers
- steals
- minutes played

How can we use this data (on a per game or per season basis) to break down what makes an NBA team perform well or poorly?

Effective Field Goal Percentage

Many coaches and players currently evaluate their shooting by looking at their field goal percentage. For example, suppose in a Dallas Mavericks–New York Knicks game the Mavericks make 45 out of 100 field goals (shoot 45%). The Knicks make 50 out of 100 field goals (shoot 50%). At first glance the Knicks shot better than the Mavericks. Suppose, however, the Mavericks shot 15 for 20 on three-pointers and the Knicks shot 1 for 5 on three-pointers. On the same number of shots the Mavericks scored 105 points and the Knicks scored 101 points. This indicates that the Mavericks

actually shot better from the field than the Knicks did. To capture this phenomenon, NBA statistical geeks have created Effective Field Goal percentage (EFG).

$$\text{EFG} = \frac{\text{(all field goals made)} + 0.5 \text{ (3-point field goals made)}}{\text{(all field goal attempts)}}.$$

In essence, EFG gives 50% more credit for making a three-pointer because a three-pointer is worth 50% more points than a two-point field goal. For our example, the Mavericks' $\text{EFG} = (45 + 0.5 \times (15))/100 = 52.5\%$ and the Knicks' $\text{EFG} = (50 + 0.5 \times (1))/100 = 50.5\%$. This example shows that EFG does a better job of capturing the quality of a team's shooting than does the traditional metric of Field Goal Percentage.

How Can We Evaluate Team Rebounding?

Raw rebounds for a team or player can be misleading. What really matters is the percentage of rebounds a team gets when they are on offense (called Offensive Rebounding Percentage, or ORP) and the percentage of rebounds a team gets when they are on defense (called Defensive Rebounding Percentage, or DRP).

Dean Oliver, a statistical consultant for the Denver Nuggets, describes a four-factor model that can be used to analyze a team's performance and to better understand a team's strengths and weaknesses.¹ The four factors (measured for both the team's offense and the team's defense) are explained below.

Four Factors for Team Offense

1. Effective Field Goal Percentage (EFG).
2. Turnovers Committed per Possession (TPP). A possession starts when a team gets the basketball and ends when they give up control of the basketball. During the last five seasons NBA teams have averaged around 92 possessions per game.²
3. Offensive Rebounding Percentage (ORP): the percentage of rebounds a team gets of their missed shots.
4. Free Throw Rate (FTR): foul shots made divided by field goal attempts. FTR is impacted by how often a team gets to the foul line as well as by their free throw percentage.

¹ Oliver, *Basketball on Paper*.

² See Kubatko et al., "A Starting Point for Analyzing Basketball Statistics."

Four Factors for Team Defense

1. Opponent’s Effective Field Goal Percentage (OEFG).
2. Defensive Turnovers Caused per Possession (DTPP).
3. Defensive Rebounding Percentage (DRP): percentage of rebounds a team gets of their opponent’s missed shots.
4. Opponent’s Free Throw Rate (OFTR): foul shots made by the opposing team divided by field goal attempts made by the opposing team.

Figure 28.1 shows the values of the four factors for teams during the 2006–7 NBA season. Figure 28.2 shows each team’s ranking for each of the four factors as well as the number of games won by each team.

	J	K	L	M	N	O	P	Q	R	S
3		Mean	0.496166667	0.496166667	0.159	0.158866667	0.270866667	0.729066667	0.246366667	0.246633333
4	Team	Wins	Offensive shooting	Defensive shooting	Offensive TOs	Defensive TOs	Offensive rebounding	Defensive rebounding	Offensive FTs	Defensive FTs
5	76ers	35	0.48	0.501	0.163	0.166	0.272	0.708	0.255	0.221
6	Bobcats	33	0.48	0.5	0.155	0.165	0.264	0.716	0.236	0.279
7	Bucks	28	0.504	0.522	0.158	0.164	0.276	0.681	0.209	0.234
8	Bulls	49	0.493	0.473	0.166	0.18	0.286	0.743	0.229	0.252
9	Cavaliers	50	0.484	0.48	0.152	0.163	0.297	0.758	0.223	0.243
10	Celtics	24	0.479	0.502	0.172	0.161	0.27	0.738	0.259	0.28
11	Clippers	40	0.481	0.488	0.162	0.148	0.272	0.747	0.28	0.249
12	Grizzlies	22	0.504	0.529	0.168	0.157	0.259	0.711	0.285	0.237
13	Hawks	30	0.471	0.503	0.17	0.163	0.292	0.709	0.263	0.268
14	Heat	44	0.506	0.485	0.159	0.154	0.249	0.733	0.222	0.232
15	Hornets	39	0.479	0.499	0.154	0.146	0.291	0.747	0.215	0.212
16	Jazz	51	0.502	0.496	0.164	0.159	0.317	0.751	0.283	0.314
17	Kings	32.4	0.491	0.513	0.149	0.166	0.231	0.725	0.289	0.24
18	Knicks	33	0.494	0.504	0.181	0.145	0.31	0.74	0.27	0.246
19	Lakers	42	0.511	0.5	0.159	0.151	0.261	0.723	0.249	0.262
20	Magic	40	0.5	0.48	0.183	0.162	0.293	0.737	0.276	0.286
21	Mavericks	67	0.509	0.477	0.151	0.157	0.287	0.75	0.256	0.265
22	Nets	41	0.504	0.49	0.157	0.152	0.246	0.744	0.245	0.266
23	Nuggets	45	0.501	0.499	0.164	0.162	0.289	0.718	0.268	0.203
24	Pacers	35	0.474	0.491	0.17	0.165	0.284	0.727	0.246	0.271
25	Pistons	53	0.488	0.477	0.135	0.162	0.283	0.709	0.237	0.234
26	Raptors	47	0.504	0.503	0.142	0.159	0.222	0.745	0.239	0.219
27	Rockets	52	0.499	0.466	0.152	0.15	0.257	0.77	0.22	0.23
28	Sonics	31	0.499	0.515	0.162	0.158	0.278	0.709	0.228	0.243
29	Spurs	58	0.521	0.471	0.15	0.155	0.242	0.757	0.235	0.201
30	Suns	61	0.551	0.492	0.147	0.152	0.227	0.719	0.215	0.206
31	Trail-blazers	32	0.483	0.508	0.163	0.15	0.282	0.73	0.241	0.267
32	Warriors	42	0.512	0.506	0.157	0.182	0.256	0.696	0.215	0.264
33	Wizards	41	0.491	0.517	0.141	0.162	0.281	0.71	0.272	0.249
34	Timber-wolves	32.4	0.49	0.498	0.164	0.15	0.252	0.721	0.231	0.226

Figure 28.1. Four-factor analysis of the NBA 2006–7 season. Source: <http://www.basketball-reference.com>.

	J	K	T	U	V	W	X	Y	Z	AA
3										
4	Team	Wins	Offensive shooting	Defensive shooting	Offensive TOs	Defensive TOs	Offensive rebounding	Defensive rebounding	Offensive FTs	Defensive FTs
5	76ers	35	25	19	19	3	16	28	12	6
6	Bobcats	33	25	17	11	5	19	22	19	27
7	Bucks	28	7	29	14	7	15	30	30	10
8	Bulls	49	17	3	24	2	9	10	22	19
9	Cavaliers	50	22	6	8	8	3	2	24	14
10	Celtics	24	27	20	28	14	18	12	10	28
11	Clippers	40	24	9	17	28	16	6	4	17
12	Grizzlies	22	7	30	25	18	21	23	2	12
13	Hawks	30	30	21	26	8	5	25	9	25
14	Heat	44	6	8	15	21	25	14	25	9
15	Hornets	39	27	15	10	29	6	6	27	4
16	Jazz	51	11	13	21	15	1	4	3	30
17	Kings	32.4	18	26	5	3	28	17	1	13
18	Knicks	33	16	23	29	30	2	11	7	16
19	Lakers	42	4	17	15	24	20	18	13	20
20	Magic	40	13	6	30	10	4	13	5	29
21	Mavericks	67	5	4	7	18	8	5	11	22
22	Nets	41	7	10	12	22	26	9	15	23
23	Nuggets	45	12	15	21	10	7	21	8	2
24	Pacers	35	29	11	26	5	10	16	14	26
25	Pistons	53	21	4	1	10	11	25	18	10
26	Raptors	47	7	21	3	15	30	8	17	5
27	Rockets	52	14	1	8	25	22	1	26	8
28	Sonics	31	14	27	17	17	14	25	23	14
29	Spurs	58	2	2	6	20	27	3	20	1
30	Suns	61	1	12	4	22	29	20	27	3
31	Trail-blazers	32	23	25	19	25	12	15	16	24
32	Warriors	42	3	24	12	1	23	29	27	21
33	Wizards	41	18	28	2	10	13	24	6	17
34	Timber-wolves	32.4	20	14	21	25	24	19	21	7

Figure 28.2. NBA team rankings by four factors, 2006–7. Source: <http://www.basketball-reference.com>.

For example, the Mavericks had an EFG of 50.9% and held their opponents to an EFG of 47.7%. The Mavs committed 0.151 turnovers per possession and caused 0.157 turnovers per possession. The Mavs rebounded 28.7% of their missed shots and 75% of their opponents' shots (or, equivalently, their opponents rebounded 25% of their missed shots). The Mavs made 0.256 free throws per FGA while their opponents made 0.265 free throws per FGA. This shows the Mavs bested their opponents on all factors except free throws. This is not surprising because the Mavs are primarily a jump shooting team that does not often drive to the basket.

The Four Factors Are Virtually Uncorrelated

The interesting thing about the four factors is that there is little correlation among them (See chapter 5 for an explanation of correlation). Recall that correlations are always between -1 and $+1$ with a correlation near $+1$ for two quantities x and y indicating that when x is big y tends to be big and a correlation near -1 indicating that when x is big y tends to be small. Figure 28.3 shows the correlations found using the Correlation option from Excel's Data Analysis Toolpak.

Notice that most of the correlations are near 0. For example, the correlation between offensive shooting and defensive turnovers is $-.10$. This indicates that if a team is better than average on offensive shooting, they will be slightly worse than average in causing defensive turnovers. Let's examine the three largest (in absolute value) correlations shown in figure 28.3.

- There is a $-.67$ correlation between defensive shooting percentage and defensive rebounding. This means that teams that give up a high shooting percentage tend to be poor defensive rebounding teams. This is reasonable because if a team fails to rebound its opponents' missed shots, they will in all likelihood get many easy inside shots or dunks on follow-up shots.
- There is a $-.47$ correlation between offensive shooting and offensive rebounding. This means that good shooting teams tend to be poor offensive

	A	B	C	D	E	F	G	H	I
1		Offensive shooting	Defensive shooting	Offensive TOs	Defensive TOs	Offensive rebounding	Defensive rebounding	Offensive FTs	Defensive FTs
2	Offensive shooting	1	-0.10823119	-0.27223683	-0.1032321	-0.471982212	-0.000659146	-0.242918	-0.31165359
3	Defensive shooting	-0.10823119	1	0.12781981	0.05038423	-0.047695316	-0.673817819	0.2461414	0.04011147
4	Offensive TOs	-0.272236826	0.127819806	1	-0.0240731	0.455388188	0.003098699	0.3435858	0.41230448
5	Defensive TOs	-0.103232065	0.050384229	-0.02407314	1	0.048171018	-0.397712659	-0.05845	0.22520563
6	Offensive rebounding	-0.471982212	-0.047695316	0.45538819	0.04817102	1	0.057779622	0.2528353	0.44674552
7	Defensive rebounding	-0.000659146	-0.673817819	0.0030987	-0.3977127	0.057779622	1	0.0662925	0.052547
8	Offensive FTs	-0.24291806	0.246141366	0.34358582	-0.0584503	0.252835284	0.066292546	1	0.36660448
9	Defensive FTs	-0.311653593	0.040111475	0.41230448	0.22520563	0.446745517	0.052547005	0.3666045	1

Figure 28.3. Correlations among the four factors.

rebounding teams. The Phoenix Suns (first in offensive shooting and twenty-ninth in offensive rebounding) are an illustration of this phenomenon. Perhaps teams loaded with good shooters do not have the tough guys needed to pound the offensive boards.

- There is a .46 correlation between offensive rebounding and offensive turnovers. This means teams that are good at offensive rebounding also tend to turn the ball over a lot. This is reasonable because good offensive rebounders tend to be poor ball handlers.

Different Paths to Team Success (or Failure)

We can use figure 28.2 to quickly zero in on the keys to success (or failure) for an NBA team. Consider the following examples:

- The 2007 champion San Antonio Spurs won because of their great shooting, holding their opponents to poor shooting opportunities, allowing few turnovers, and by holding down fouls and giving up few points from the foul line.
- The 2007 Phoenix Suns were successful because of great shooting, lack of turnovers, and giving up few points from the foul line. The Suns succeeded despite their poor offensive and defensive rebounding and lack of free throws.
- The 2007 Memphis Grizzlies had the league's worst record. Their shooting percentage defense was poor, they allowed lots of turnovers, and they rebounded poorly.
- The 2007 New York Knicks had a 33–49 record, due almost exclusively to lots of turnovers committed on offense, while their defense caused very few turnovers.

How Important Are the Four Factors?

Can we estimate the relative importance of the four factors? Recall that in chapter 18 we used regression to predict NFL team performance based on measures of passing and rushing efficiency, and of turnover frequency. (See chapter 3 for a description of regression analysis.) We can use regression in a similar way to evaluate the importance of the four factors in basketball performance. Running a regression using the data in figure 28.1 can predict a team's number of wins from the following four independent variables:

- EFG – OEFG
- TPP – DTPP

- ORP – DRP
- FTR – OFTR.

The results of the regression are shown in figure 28.4. From these results we can predict

$$\begin{aligned} \text{games won} = & 41.06 + 351.88(\text{EFG} - \text{OEFG}) \\ & + 333.06(\text{TPP} - \text{DTPP}) + 130.61(\text{ORP} - \text{DRP}) \\ & + 44.43(\text{FTR} - \text{OFTR}). \end{aligned}$$

	A	B	C	D	E
1	SUMMARY OUTPUT				
3	Regression Statistics				
4	Multiple R	0.953165405			
5	R Square	0.908524289			
6	Adjusted R Square	0.893888175			
7	Standard Error	3.533831698			
8	Observations	30			
9					
10	ANOVA				
11		df	SS	MS	F
12	Regression	4	3100.719505	775.17988	62.0741478
13	Residual	25	312.1991617	12.487966	
14	Total	29	3412.918667		
16		Coefficients	Standard Error	t Stat	P-value
17	Intercept.	41.05829628	0.645276505	63.628996	3.54259E-29
18	Shooting Dev.	351.8800481	28.36841836	12.403936	3.52088E-12
19	Turnover Dev.	333.0598616	52.17592394	6.3834013	1.10625E-06
20	Rebound Dev.	130.6051555	22.96495865	5.6871496	6.37464E-06
21	Free Throw Dev.	44.42983641	23.55911647	1.8858872	0.070980928

Figure 28.4. Four-factor regression.

These four independent variables explain 91% of the variation in the number of games won. The standard error of 3.53 means we are 95% sure our predicted wins will be within $2(3.53) = 7.06$ wins of the actual number of wins.

To measure the impact of the four factors on wins, we can look at the correlations between the four factors and wins:

- EFG – OEFG has a .85 correlation with wins and by itself explains 71% of the variation in wins.
- TPP – DTPP has a 0.38 correlation with wins and by itself explains 15% of variation in wins.
- ORP – DRP has a 0.25 correlation with wins and by itself explains 6% of variation in wins.

- FTR – OFTR has a -0.01 correlation with wins and by itself explains virtually none of the variation in wins.

This analysis indicates that NBA teams' differential on shooting percentage is by far the most important factor in their success.

The relative importance of the four factors are summarized as follows:

- A 0.01 improvement in EFG – OEFG is worth 3.5 wins. That is any of the following improvements:
 1. improve our EFG 1% (say, from 47% to 48%);
 2. reduce our opponent's EFG by 1%; or
 3. improve our EFG by 0.5% and cut our opponent's EFG by 0.5% should on average cause us to win 3.5 more games.
- A 0.01 improvement in TPP – DPPP is worth 3.3 wins. Thus any of the following improvements:
 1. one less turnover per 100 possessions; or
 2. one less turnover per 200 possessions and causing one more turnover per 200 possessions would lead to 3.3 more wins.
- An increase of 0.01 in ORP – DRP would lead on average to 1.3 more wins per season. Thus any of the following combinations would be expected to lead to 1.3 more wins:
 1. one more offensive rebound per 100 missed shots;
 2. one more defensive rebound per 100 shots missed by opponent; or
 3. one more offensive rebound per 200 missed shots and one more offensive rebound per 200 shots missed by an opponent.
- An increase of 0.01 in FTR – OFTR would be expected to lead to 0.44 wins. Therefore any of these three combinations would be expected to lead to 0.44 wins:
 1. one more free throw made per 100 field goal attempts;
 2. one less free throw given up per 100 field goal attempts by opponent; or
 3. one more free throw made per 200 field goal attempts and one less free throw given up per 200 field goal attempts.

In summary, Dean Oliver's decomposition of a team's ability into four factors provides a quick and effective way to diagnose a team's strengths and weaknesses. Of course, the four-factor model may be applied to final or season-in-progress cumulative team data, or simply to the box score data for any particular game.