

Journal of Quantitative Analysis in Sports

Manuscript 1355

Are the “Four Factors” Indicators of One Factor? An Application of Structural Equation Modeling Methodology to NBA Data in Prediction of Winning Percentage

Tarek Baghal, *University of Nebraska–Lincoln*

Are the “Four Factors” Indicators of One Factor? An Application of Structural Equation Modeling Methodology to NBA Data in Prediction of Winning Percentage

Tarek Baghal

Abstract

Significant work has gone into the development of team and individual statistics in the NBA; for example, the team measures of the “Four Factors.” Less work has been conducted using multivariate analyses of these metrics, including identifying possible new statistical techniques to analyze these data. In particular, this research examines the feasibility of using structural equation modeling (SEM) for multivariate analyses of NBA Four Factors data. SEM consists of both confirmatory factor analysis (CFA) and path modeling. Before SEM is employed, this research first examines the effects of offensive and defensive Four Factors in a linear regression model, expanding previous research and providing a baseline for the SEM. In doing so, the data show the importance of effective field goal percentage. Next, structural equation modeling is employed. The CFA finds that offensive Four Factors are indicators of a single latent factor, labeled “offensive quality.” The “defensive quality” latent factor is estimable when replacing opposing teams’ free throw rate with steals per possession. The SEM is extended to regress winning percentage on latent offensive and defensive quality as well as salary. Salary is an important and often overlooked part of multivariate models examining team statistics, but it is easily incorporated in SEM. The explained variance for the regression in the SEM is similar to that of the linear regression model and indicates the importance of both offensive and defensive quality, with offensive quality having a larger effect. Team salaries are related to offensive quality, but not defensive quality or winning. As such, a second structural equation model is proposed where the effect of salary on winning is mediated by its relationship with offensive and defensive quality. Since salary is related to offensive quality but not defensive quality, and offensive quality is more important to winning percentage, this suggests that money spent is done so for offensive performance and affects winning through the performance paid for. These results suggest potential team strategies, as well as the applicability of SEM to sports analytics, not only to NBA data, but to other sports data as well.

KEYWORDS: NBA, Four Factors, salary, structural equation modeling

Author Notes: The author would like to thank Nikki Siddiqi, Rene Bautista, Jim Bovaird, Justin Mills, Ibad Siddiqi, Mathew Stange, Aaron D. Yates, and an anonymous reviewer for helpful comments in developing this research.

1 Introduction

The desired outcome for team performance in National Basketball Association (NBA) games, like in other sports, is winning. Many variables may affect game outcomes, and statistical analyses allow for an examination of which variables are of greatest importance. In turn, such analyses could be used in creating team strategies, as well as direction for efficient team spending on player salaries. Research on the NBA in this regard is in the relatively nascent stages, with recent work appearing in order to create a common framework for analyses of NBA data (Kubatko et al. 2007). This work has focused on the development of metrics to assess team and player performance on a variety of offensive and defensive dimensions.

Of particular interest in team performance analyses are the metrics the “Four Factors”: effective field goal percentage (EFG), free throw rate (FTR), turnovers per possession (TPP), and offensive rebounding percentage (ORP). These can be calculated for both offensive and defensive aspects of a team. These metrics are of importance and selected for analysis for two major reasons. First, these measures capture the components that are thought to make up the overall offensive or defensive performance (Kubatko et al. 2007). Second, research has shown that the Four Factors are strongly related to winning (Teramoto and Cross 2010).

Analyses of these metrics have increased the understanding of the importance of these factors overall, and the relative importance of each factors. Research shows that the mean of the Four Factors vary over time, with a general overall decline in offensive responding percentage and apparently turnovers as well (Kubatko et al. 2007). Other work has attempted to place weights on each of the Four Factors, in order to show the relative importance in regards to team winning (Küpfer 2005). This work suggests that effective field goal percentage is most important, followed by turnovers per possession, with offensive rebounding percentage and free throw rate being relatively less important (Küpfer 2005).

Multivariate analyses could further the findings using the Four Factors to improve understanding of how it affects game outcomes. However, multivariate analyses of these and their relation to winning has been somewhat limited. Past research has examined team statistics for these metrics and attempted to relate them to winning, or have developed projection systems based on individual player projections (e.g. SCHOENE) (Doolittle and Pelton 2010).

Recently, multivariate models (i.e. ordinary least squares) were used in estimating the relative impact of team offensive and defensive Four Factors in winning percentage in the regular season and playoffs (Teramoto and Cross 2010). Examination of the standardized coefficients found that offensive and defensive EFG were the most important factors in determining winning in both

the regular season and playoffs, with defensive EFG apparently more important in winning playoff series. Offensive and defensive FTR, while significant predictors of winning percentage, were relatively the least important of the Four Factors (Teramoto and Cross 2010). The findings generally confirmed past arguments about the relative importance of the factors (e.g. Küpfer 2005)

Although such research is a step in the right direction, there are some limitations. The first limitation is the lack of cost data (i.e. team salary) incorporated into the multivariate analyses, which fails to account for the effect of spending on winning. Including cost allows examination of potential efficiencies (or inefficiencies) in spending as it relates to winning. For example, if defensive statistics are more strongly related to winning than offensive statistics, but salary is unrelated to defensive statistics and strongly related to offensive statistics, this denotes a possible spending inefficiency.

A second issue is that each of the metrics of the Four Factors is likely related to the same construct. Some have argued that the Four Factors reflect components of the offensive and defensive ratings, a metric based on points per possession (Kubatko et al. 2007). However, this suggests that the Four Factors are causal to the offensive rating; it does not answer what causes the Four Factors. It is possible that the Four Factors are indeed related to and caused by the same underlying construct of interest, team offensive and defensive quality. The Four Factors may be driven by the underlying quality of a team, and any single measure is limited in terms of being an imperfect (or incomplete) indicator of offensive and defensive quality. Thus, the effect of effective field goal percentage on winning tells a limited story, as EFG is only one component of an offense leading to points per possession.

Therefore, it is important to estimate the underlying construct of interest. There are several ways that this may be done; with numeric measures such as the Four Factors, the most promising may be structural equation modeling (SEM) (Kline 2005). SEM estimates the “true” level of a latent construct (e.g. offensive or defensive “quality”) based on several indicators using confirmatory factor analysis (i.e. the measurement portion of the model). In addition, SEM allows for estimation of the relationships between variables of interest, as one would in a regression model, by incorporating path analyses between the variables and estimated factors (i.e. the structural portion of the model) (Kline 2005). Using this method, it is possible to estimate the underlying quality measures of interest and relate them to salary and winning.

There were several objectives to the current research. The first was to conduct regression analyses on winning and the Four Factors using a larger number of years than Teramoto and Cross (2010), and compare results. Second was to use SEM to determine if the Four Factors could be modeled as indicators of single latent variables for offensive and defensive quality. If this could be done,

the third objective was to relate the identified latent quality variables to winning percentage. Additionally, team salaries were incorporated into the SEM models in order to estimate the relationship of cost with offensive and defensive quality as well as winning percentage. Further, this research had the goal of introducing SEM as an effective statistical tool in sports analytics.

2 Data and Methods

Annual team data were obtained from Basketball-Reference.com (2010) between the 1995-1996 and 2008-2009 seasons. The start date was selected as it coincided with the first year of play for the Vancouver Grizzlies and Toronto Raptors, increasing the number of teams in the league to 29. In 2004, the Charlotte Bobcats joined the league. Treating each year as an independent observation leads to a total sample size of 411. Salary data came from Patricia Bender’s collection of NBA data (Bender 2010).

Team possessions and Four Factors statistics were calculated using formulas commonly accepted by other NBA researchers (Kubatko et al. 2007). In particular, these are:

$$(1) EFG = (FGM + 0.5 \times 3PM)/FGA.$$

$$(2) FTR = FTM/FGA$$

$$(3) TPP = TO_t/POSS_t^1$$

$$(4) ORP = OREB_t/(OREB_t + DREB_o)$$

Where FGM is field goal made, FGA is field goal attempted, 3PM is three point field goals made, FTM is free throw made, TO is turnovers, OREB is offensive rebounding, DREB is defensive rebounds, and POSS is possessions. Subscript t indicates team of interest, subscript o indicates opposing team.

For use in the SEM, the Four Factors were all multiplied by 100 to indicate percentage rather than proportion for EFG, DEFG, ORP, and DORP, number of turnovers per 100 possessions, and number of free throws made per 100 field goal attempts. This transformation makes interpretation easier and consistent with other measures, such as offensive and defensive ratings (Kubatko et al. 2007). The mean of team salaries increases rapidly with each passing year, e.g. NBA mean salary in 1995-1996 = \$24.97 million, mean salary in 2008-2009 = \$71.66 million. Therefore, these data were transformed to a ratio of annual team salary for that year and NBA mean salary (i.e. salary of team in year/NBA overall mean in year). The result allows for a consistent metric across all years, interpretable as the relative size of team salary compared to the entire league, with

¹ $POSS = FGA_t + 0.44 \times FTA_t - OREB_t + TO_t$

a ratio of 1 indicating team salary for a given year was equal to the NBA overall mean salary for that year. For example, in 2008-2009 the New York Knicks had a team salary about 1.35 times larger than the league average of \$71.66 million. All structural equation models were estimated using Mplus (Muthen and Muthen 2009).

3 Results

The means and standard deviations of the offensive Four Factors years across all teams and years are presented in Table 1. As the means across all teams and years for offensive and defensive Four Factors are calculated as such to be nearly identical, only offensive Four Factor means are presented. As can be seen, when taking three point shots into account, teams approach fifty percent effectiveness in field goal attempts. One way to interpret the mean FTR is that for each shot a team takes, on average it makes 0.23 free throws (i.e. 10 attempts yields slightly more than 2 made free throws). Slightly more than 15% of possession leads to turnovers, while teams get about 29% of the rebounds from their own missed shots.

Table 1: Offensive and Defensive Four Factors Means, All Teams and Years

<i>Variable</i>	<i>Mean (SD)</i>
EFG	0.484 (0.021)
FTR	0.237 (0.026)
TPP	0.156 (0.014)
ORP	0.287 (0.028)

EFG = Effective Field Goal %, FTR = Free Throw Rate, TPP = Turnovers Per Possession, ORP= Offensive Rebound Percent

Previous research has regressed winning percentage on these offensive and defensive Four Factors, i.e. effective field goal percentage (EFG), free throw rate (FTR), turnovers per possession (TPP), and offensive rebounding percentage (ORP) (Teramoto and Cross 2010). This was replicated as a first step in the current research. However, the number of years, and thus number of cases, was increased in the current research from 10 to 14 (with the sample increasing from 295 to 411). Increasing the years and sample size serves as a check on past results

while increasing certainty and how applicable these are over time. Results of the current ordinary least squares (OLS) regression of winning percentage on offensive and defensive Four Factors are presented in Table 2 (defensive counterparts of the offensive Four Factors are preceded by D in the abbreviation). A White test of the residuals (White 1980) suggested that heteroskedasticity was not an issue ($\chi^2_{44} = 49.50$, $p = .2631$), and thus unadjusted standard errors are reported.

Table 2: Regression of Winning Percent on Offensive and Defensive Four Factors

<i>Variable</i>	<i>Coefficient (SE)</i>	<i>Standardized Coefficient</i>
EFG	4.740* (0.160)	0.620
FTR	0.782*(0.131)	0.125
TPP	-4.410* (0.280)	-0.373
ORP	1.564*(0.130)	0.267
DEFG	-3.666* (0.163)	-0.450
DFTR	-0.875* (0.128)	-0.155
DTPP	3.274* (0.270)	0.280
DORP	-1.666* (0.158)	-0.249
Constant	0.196 (0.119)	0
R ² =0.859, N=411, *p<0.01		
<i>EFG = Effective Field Goal %, FTR = Free Throw Rate, TPP = Turnovers Per Possession, ORP= Offensive Rebound Percent, D indicates Defensive aspects of measures</i>		

The results are similar to past results using a more restricted data set (Teramoto and Cross 2010), with all coefficients significant and in the expected direction. Although all of the offensive and defensive Four Factors are important in predicting winning percentage, the importance of effective field goal percentage is apparent. Unlike the past work in Teramoto and Cross, who found that offensive and defensive EFG are nearly the same in the effect on regular season winning, this analysis shows that offensive field goal percentage carries more importance in predicting winning percentage. A contrast between absolute values of offensive and defensive effective field goal percentage shows that the coefficient for offensive EFG is significantly larger, $F(1,402) = 25.47$, $p < 0.01$. The second most important measure was TPP, followed by DTPP, ORP, and DORP, with defensive and offensive free throw rate, respectively, found to be relatively the least important of the Four Factors.

As previously noted, each of the offensive and defensive Four Factors are conceptually linked to the overall offensive and defensive performance of the team. While the above analysis shows the relative importance of each of the

individual metrics, it is possible that these metrics are also only observed outcomes of overall offensive and defensive quality. Considering a single measure such as EFG may be limited because it is only one aspect of offensive or defensive quality. This quality is latent, directly unobservable, only manifested through measures such as the Four Factors. Therefore, each of the offensive (defensive) Four Factors is driven by the level of offensive (defensive) quality.

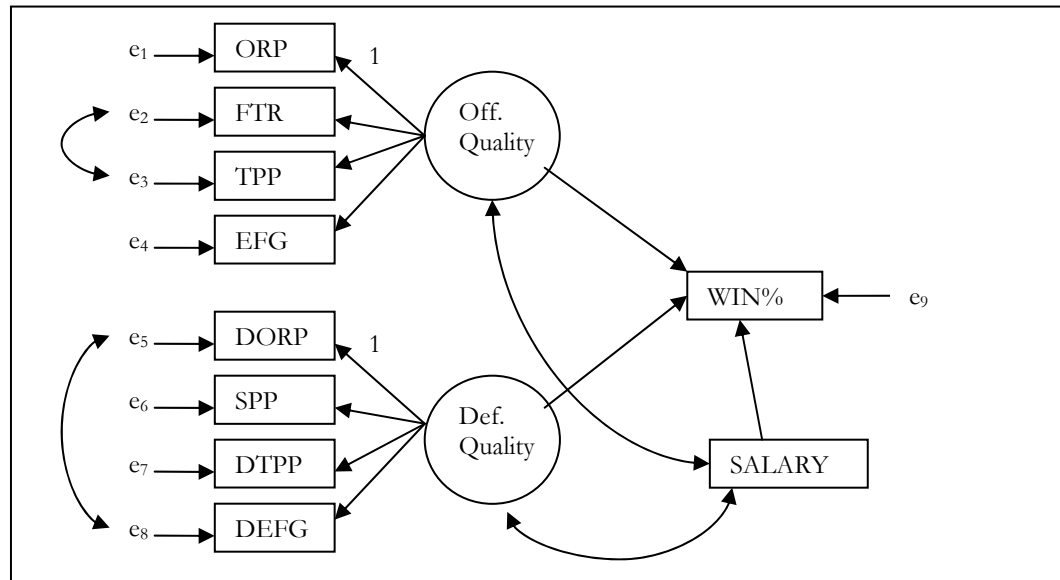
To test whether these Four Factors are indeed indicators of latent characteristics, as well as examine the effects of these latent traits on winning, confirmatory factor analysis (CFA) was conducted. The offensive set of Four Factors loaded onto a single factor, indicating these all measure one construct, labeled “offensive quality”. However, the defensive Four Factors failed to load onto one factor; further analysis indicated this was due to the defensive FTR measure. This seems plausible, as free throws made by the other team are not completely in the control of the defense, only the number of free throws attempted. Other measures of defensive quality were identified in order to keep equal number of indicators for both the latent constructs. Steals per possession (SPP) ($M=0.083$, $SD=0.010$) was selected as it is a defensive statistic that is within the control of the defense.

The standard OLS regression model reported in Table 2 was largely unchanged when SPP replaced DFTR ($R^2=0.844$), with the standardized beta for SPP = 0.065 ($p<.10$), somewhat smaller in effect than defensive free throw rate. In addition, when SPP replaced defensive FTR in the SEM, a single latent variable was identified, labeled “defensive quality”. For both offensive and defensive latent factors, the loading for ORP and DORP was constrained to 1, as one observed variable (or latent variance) must be constrained to have an estimable model (Kline 2005). Selecting observable variables as those constrained is consistent with the Unit Loading Identification (ULI) frequently used for model estimation (Kline 2005).

Modification indices output by the program indicated that allowing residuals of offensive FTR to covary with the residuals of TPP, as well as allowing the residual of defensive ORP to covary with defensive EFG residual would improve model fit. Since it is conceivable logical that both of these covariations could occur, the model included these covariations. That is, it seems reasonable that the amount of free throws made in relation to number of shots taken is related to turnovers (which may also affect field goals attempted, part of FTR). Similarly, how many offensive rebounds the other team gets may affect that team’s field goal percentage (e.g. shots closer to/farther from the basket). Salary (annual salary/league mean salary ratio) entered the model first as a correlated variable to the latent offensive and defensive factors. Next, winning percentage was regressed on offensive and defensive “quality” latent factors as well as salary. Using common convention for the diagramming of SEM, Figure 1

illustrates the final model described here, where boxes represent observed variables, circles represent latent variables, the number one represents the ULI constraint, and e indicates the residuals for a given endogenous variable.

Figure 1. SEM of Winning Percent on Offensive and Defensive Quality Latent Variables



EFG = Effective Field Goal %, FTR = Free Throw Rate, TPP = Turnovers Per Possession, ORP= Offensive Rebound Percent, D indicates Defensive aspects of measures, SPP= Steals Per Possession

This model was estimable, although many of the model quality indicators did not meet standard levels (Hu and Bentler 1999): model $\chi^2_{29} = 720.410$ $p < .001$, CFI=0.578, RMSEA = 0.241, 90%CI RMSEA = 0.226, 0.256, SMSR = 0.118. Parameter and variance estimates are presented in Table 2. Panel A presents the measurement portion of the model (the estimation of the latent factors) with Panel B displaying the results of the structural portion of the model (the path/regression model).

Turning to the measurement portion of the model, it is important to note that based on the loadings, the scaling of the offensive and defensive quality latent indicators is similar, and thus interpretable in a similar metric. The parameters' paths from the latent variables to the indicators are interpreted as in a regression, with an increasing of one unit in the latent variable leading to an expected change in the indicator variable.

Generally, all parameters in the measurement model are in the expected directions, e.g. offensive quality is directly related to EFG and defensive quality inversely related to DEFG. The exceptions may be that higher ORP and DORP

(see standardized coefficients) are related to lower offensive and defensive quality, respectively. This may be for a number of factors, possibly related to the shot previous or following the rebound. For example, if most offensive rebounds come from long shots with rebounds being further the basket, follow-up shots may not be as optimal as expected. It also may be simply related to the current model specification, where ORP and DORP loadings are constrained to unity. However, when this constraint is removed, and the unity constraint is placed on the latent variance (UVI, see Kline 2005), the same pattern emerges.

Table 3: SEM Regressing Winning Percentage on Offensive and Defensive Quality and Salary

<i>Panel A: Measurement Model</i>				
	Loading (S.E.)	Standardized Parameter	Error Variance	Factor Variance/ Disturbance
Offensive Quality				0.753
ORP	1.00 (0.00)	-0.316	6.798	
FTR	0.406* (0.169)	0.137	6.454	
TPP	-0.788* (0.145)	-0.504	1.375	
EFG	1.730* (0.292)	0.715	2.157	
<i>FTR with TPP</i>			1.274	
Defensive Quality				0.979
DORP	1.00 (0.00)	0.412	4.783	
SPP	0.940* (0.116)	0.943	0.108	
DTPP	1.146* (0.138)	0.824	0.608	
DEFG	-0.286* (0.115)	-0.143	3.812	
<i>DEFG with DORP</i>			1.152	

* indicates $p < .05$. Bold indicates latent variables.

EFG = Effective Field Goal %, FTR = Free Throw Rate, TPP = Turnovers Per Possession, ORP = Offensive Rebound Percent, D indicates Defensive aspects of measures, SPP = Steals Per Possession

<i>Panel B. Structural Model</i>			
	Coefficient (S.E.)	Standardized Parameter	R ²
Winning Percentage ON			0.820
Offensive Quality	17.694* (3.127)	0.956	
Defensive Quality	6.782* (1.188)	0.418	
Salary	0.473 (0.326)	0.067	
<i>Off. Quality with Salary</i>		0.457* (0.141)	
<i>Def. Quality with Salary</i>		-0.094 (0.117)	

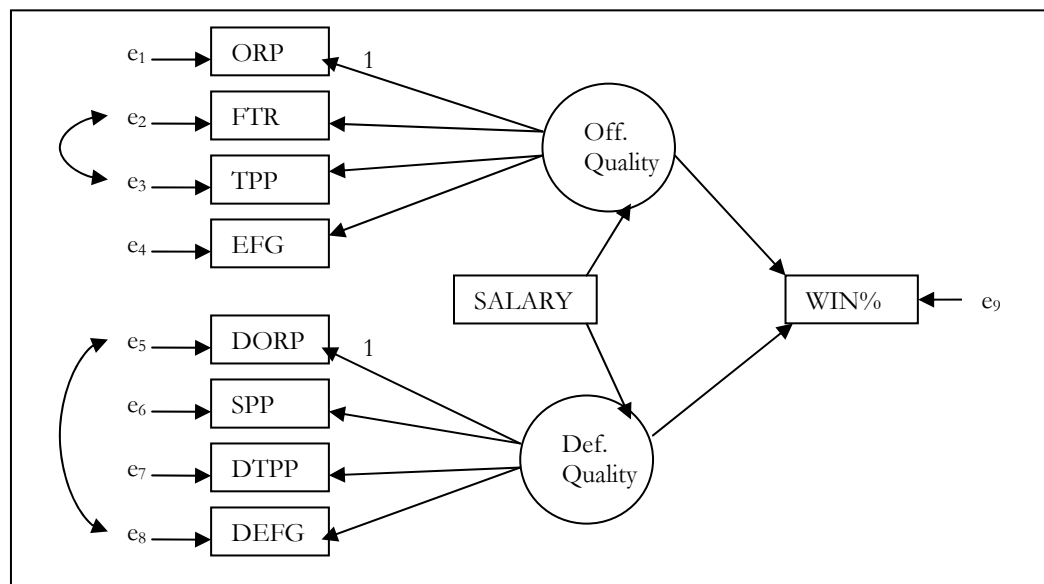
* indicates $p < .05$. Bold indicates latent variables.

Still, the other indicators follow relationships with the underlying latent variable as expected. For example, a one unit increase in offensive quality leads to an expected increase in 1.73% in effective field goal percent (EFG), and a one unit increase in defensive quality leads to an expected increase leads to an expected increase of 1.15 other team turnovers per one hundred possessions. Increases in offensive quality also leads to increases in own team free throw rate, with decreases in own team turnover per possession. Defensive quality increases also leads to increases in own team's steals per possession and decreases in other teams effective field goal percentage.

In regards to the structural portion of the model, there are clear effects of offensive and defensive quality latent factors on winning percentage. Of importance is that the amount of variance explained in winning percent ($R^2=0.820$) is approximately the same as the OLS regression discussed above ($R^2=0.859$ for the original model, $R^2=0.844$ for the model replacing DFTR with SPP). This suggests that, at the least, the explanatory power of the model is not decreased significantly by using SEM. Further, both offensive and defensive quality are significant effects, leading to large gains in winning percentage. However, offensive quality leads to greater increases in winning percentage than defensive quality, as seen in both the unstandardized and standardized coefficients. This is consistent with the OLS regressions above and those in Teramoto and Cross, which show the possible relative importance of offensive traits in the regular season.

However, unlike those regressions, the SEM show how each of the Four Factors come together to have an overall effect as opposed to examining the constituent effects. The current analysis suggests that a one unit increase in offensive quality will lead to a 17.69% increase in winning percent, and a one unit increase in defensive quality leads to an expected 6.78% increase in winning percent. This also suggests that increases in a unit in the offensive and defensive quality are not easily accomplished. Although offense quality may be relatively more important than defensive quality, at least in the regular season, it is clear that defense plays a large role winning games.

Figure 2. SEM of Winning Percent on Offensive and Defensive Quality Latent Variables, with Salary a Prior Causal Effect



EFG = Effective Field Goal %, FTR = Free Throw Rate, TPP = Turnovers Per Possession, ORP= Offensive Rebound Percent, D indicates Defensive aspects of measures, SPP= Steals Per Possession

An interesting pattern of relationships emerge between salary and other variables, as indicated in Table 3. First, salary is not significantly correlated with defensive quality ($r=-0.094$, $p=0.43$), but is correlated with offensive quality ($r=0.457$ $p<0.01$). This suggests that money is spent on offensive quality, but not necessarily on defensive quality. Additionally, the effect of salary on winning percentage is also found to be non-significant at standard levels (although the coefficient is in the expected direction, with more salary leading to increases in expected win percent). It is possible this is due to the relationship with offensive quality, which also has a larger impact on winning percentage. That is, salary has an indirect effect; salary is directly related to offensive quality (and possibly

defensive quality), which in turn helps predict winning (i.e. the effect of salary on winning is mediated through its effects on team quality measures). A benefit of SEM is that such a relationship can be modeled. This model is presented in Figure 2, and the estimated coefficients in Table 4. Since the measurement portion of the model remains unchanged, only the structural portion of the model is reported in Table 4.

Table 4: SEM Regressing Winning Percentage on Offensive and Defensive Quality with Salary as Prior Cause

<i>Structural Model</i>			
	Coefficient (S.E.)	Standardized Parameter	R ²
Offensive Quality			0.077
ON			
Salary	0.126* (0.024)	0.278	
Defensive Quality			0.001
ON			
Salary	-0.012 (0.021)	-0.030	
Winning Percentage			0.790
ON			
Offensive Quality	13.952* (0.759)	0.856	
Defensive Quality	4.498* (0.811)	0.248	

* indicates $p < .05$. Bold indicates latent variables.

As with the previous model, the relationship between salary and defensive quality is not significant. However, there is a clear relationship between salary and offensive quality, with increases in salary leading to increases in the offensive quality. Evidently, when teams spend on players, they do so in regards to offensive characteristics rather than defensive. This may be an effective spending strategy, as can also be seen from Table 4. Although the effects of offensive and defensive quality are diminished when salary is used as causally prior compared to the original model, the effects are still strong and similar in nature. Increasing offensive quality leads to greater gains than does defensive quality. Since salary is more clearly linked to offensive quality, spending on offense indirectly leads to a

greater winning percentage. However, although the model does indicate a possible scenario for predicting winning percentage, it is important to note that the explained variance of winning percentage (R^2) did decrease slightly. This decline was minor, however, and the overall R^2 was still in the range of other models (0.790).

4 Discussion

Generally, this research shows the potential use of structural equation models on NBA data, and more generally, in sports analytics. SEM appears to be a useful tool to consider, especially given the need for more multivariate analyses and in situations when many indicators of an underlying dimension are present. Although past research was expanded and replicated, it examined each offensive and defensive measures separately, rather than as a whole. The SEM models presented here allow for estimation of the relation between the latent characteristics of interest (quality) and to outcomes of interest, i.e. winning. Further, such models may be useful in analysis of other sports statistics, especially those where team data uses different metrics for different positions on offense and defense (e.g. football) may be especially open to the use of SEM. For example, rushing yards per attempt, pass yards per attempt, and yards after the catch are in many ways distinct but may all be driven by an underlying quality of the team as a whole, including coaching strategy and player skills.

Some limitations and areas of further research should be pointed out. First, although SEM appears to be a useful tool to be used in sports statistical analyses, potentially across a number of sports, the models examined in this paper may be improved. Neither model achieved overall model fit statistics generally used as standards is structural equation modeling. However, the explained variance (i.e. R^2) for the variable of interest, winning percent, was consistently high and at levels found in other analyses examining the same outcome. Next, the reasons why some measures did not fit well into the latent model, such as defensive free throw rate, or why offensive rebounding percentage and the defensive counterpart loaded in unexpected ways should be examined further. Some reasons were posited as to why in this paper, but further studies should examine these using some empirical methods.

Additional research should also focus how salary affects winning and the best way to conceptualize this effect. This can include statistical issues or those on how to use results from statistical models in decision-making. On the statistical side, it may be that salary determines and is determined by offensive quality (i.e. a circular pattern). For example, it may be that more spending brings in better players, who then based on greater incentive to perform, perform at a higher level, potentially driving up contract demands. This information could be used by teams

to better develop strategies for what types of players to spend money on. Similarly, more research is needed on the player level in addition to team level data. Structural equation models, as well as other statistical models such as multilevel models, can examine player level impacts on outcomes of interest such as offensive quality or winning. However, individual level data leads to the potential problem of nonindependence of observations. Specifically, it is difficult to ascertain how much of the player performance is due to teammates or players defending. The adage “he makes his teammates better” is a potential statistical problem, one that should be explored.

5 Conclusions

The current research produces several important findings. First, based on the SEM models, it is found that the Four Factors are indicators of more general team traits, labeled offensive and defensive quality. This suggests that the quality of the team can be measured and drives other aspects of the team, such as team efficiency and winning percentage. Interestingly, defensive free throw rate did not load on the defensive quality latent variable. It is suggested here this may be due to the fact that free throws made by the other team is only affected by the defense through the number of free throws allowed. Once the foul is committed, whether the free throw is made or not is out of the hands of the defensive team. As such, it may not be as good of an indicator of defensive quality. Once this measure was replaced by steals per possession, a statistic more in control of the defense, then a latent variable for defensive quality was estimable. Additionally, offensive rebounding percentage and its defensive counterpart both seemingly are affected by the quality latent variables in ways opposite than expected. The reasons for this are less clear, potentially related to the shot prior to or following the offensive rebound.

Second, the research indicates the potential greater importance of offense than may have been thought previously. Previous research and research using techniques such as ordinary least squares suggest that although offense may be more important in winning percentage, at least in the regular season, there is not much difference between offensive and defensive statistics. The above findings suggest that possible overall team offensive quality increases winning much more than defensive quality does. This may explain why teams perceived to have high powered offense and poor defense (e.g. Phoenix Suns) have had success in the regular season.

Third, spending on team salary seems to be mainly related to offensive quality, while directly unrelated to defensive quality and winning percentage. However, spending on team salary seems to have a moderated effect on winning, whereby spending leads to greater offensive quality, which in turn leads to greater

winning percentage. Further, since offensive quality is relatively more important to winning percentage than defensive quality, this may suggest that spending is directed to the correct team characteristics. However, since spending is unrelated to defensive quality, which does have a significant impact on winning, this possibly suggests more spending should focus on defensive characteristics, although not necessarily at the expense of offensive quality. Using the relationships between salary, quality factors, and the observed indicators, teams could also begin to examine potential characteristics to target in search of a higher winning percentage. For example, offensive quality may be best increased by improving effective field goal percentage to the team, but decreasing turnovers does not follow far behind in regards to potential importance.

References

- Basketball-Reference.com. 2011. <http://www.basketball-reference.com/>, accessed 2/28/2011
- Bender, P. 2011. "Patricia's Various NBA Basketball Stuff" <http://www.eskimo.com/~pbender/>, accessed 2/28/2011
- Doolittle, B. and Pelton, K. 2010 *Pro Basketball Prospectus 2010-2011*, 2010.
- Hu, L. and Bentler, P.M. 1999. "Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives." *Structural Equation Modeling* vol. 6, pp. 1-55
- Kline, R.B. 2005. *Principles and Practices of Structural Equation Modeling*, 2nd Edition, New York, NY
- Kubatko, J., Oliver, D. Pelton, K. and Rosenbaum, D.T. 2007 "A Starting Point for Analyzing Basketball Statistics" *Journal of Quantitative Analysis in Sports*, vol. 3, iss. 3, Article 1.
- Küpfer, E. 2005. "Team Similarity." *APBRMetrics forum*, <http://www.sonicscentral.com/apbrmetrics/viewtopic.php?p=118&sid=ff925dbe120f6c6f86e596d0d6dabbb8>, accessed 2/28/2011
- Muthén, L. K., & Muthén, B.O. (2007). *Mplus user's guide* (5th Ed.). Los Angeles, CA: Muthén & Muthén.
- Teramoto, M. and Cross, C.L. 2010. "Relative Importance of Performance Factors in Winning NBA Games in Regular Season versus Playoffs" *Journal of Quantitative Analysis in Sports*, vol. 6, iss. 3, Article 2
- White, H. 1980. "A heteroscedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroscedasticity" *Econometrica* vol. 48, pp. 817-838.