Modeling and forecasting the outcomes of NBA Basketball games

Hans Manner*[1]

[1] *Institute of Econometrics and Statistics, University of Cologne*

July 10, 2015

Abstract

This paper treats the problem of modeling and forecasting the outcomes of NBA basketball games. First, it is shown how the benchmark model in the literature can be extended to allow for heteroscedasticity and treat the estimation and testing in this framework. Second, time-variation is introduced into the model by (i) testing for structural breaks in the model and (ii) introducing a dynamic state space model for team strengths. The in-sample results based on eight seasons of NBA data provide some evidence for heteroscedasticity and a few structural breaks in team strength within seasons. However, there is no evidence for persistent time variation and therefore the *hot hand* belief cannot be confirmed. The models are used for forecasting a large number of regular season and playoff games and the common finding in the literature that it is difficult to outperform the betting market is confirmed. Nevertheless, it turns out that a forecast combination of model based forecasts with betting odds can outperform either approach individually in some situations.

Keywords: Sports forecasting, paired comparisons, NBA basketball data, heteroscedasticity, time-variation

*manner@statistik.uni-koeln.de

1 Introduction

The statistical modeling of sports data has become a large topic of research over the past decades. Detailed data of high quality has become easily available due to its publication and distribution via the internet, which allows researchers to address a variety of questions. One problem of particular interest is the prediction of the outcomes, both in terms of the final score and the winning team; see Steckler et al. (2010) for an overview. This is closely related to the issue of modeling the strength of each player or team involved in the competition of interest. The best known example of such an approach is the Elo rating in chess (Elo 1978), but similar statistical methods have been applied in many different sports. Such a strength, or rating, can be obtained by variations on the statistical method of paired comparison models by Bradley and Terry (1952) and David (1959). A notable methodological innovation was the introduction of dynamic models of paired comparison in Glickman (1993) and Fahrmeir and Tutz (1994). This approach has been applied to soccer (Fahrmeir and Tutz 1994 or Koopman and Lit 2014), chess and tennis (Glickman 1999), and football (Glickman 2001, Glickman and Stern 1998), finding evidence of time-varying team/player ratings.

The present paper treats the modeling and prediction of national basketball association (NBA) basketball games. The NBA is the most important and strongest professional basketball league in the world, consisting of 30 teams/franchises. With revenues of 4.6 billion US$ and an average team worth of 634 million US$ the league has a high economic relevance.

Statistical models for various aspect of basketball have been suggested in the literature. Early contributions introducing the regression based approach to Basketball modeling are Stefani (1977a) and Stefani (1977b). The National Collegiate Athletic Association (NCAA) basketball tournament has been analyzed and modeled in several studies, e.g., Schwertman et al. (1991), Carlin (1996) or Harville (2003), with a focus on computing win probabilities and accurate team rankings. A further topic that is often addressed in the literature is the home court advantage, studied in Harville and Smith (1994), Jones (2007, 2008), or Entine and Small (2008). Other studies focus more on the relevance of game statistics, such as Kubatko et al. (2007) who introduce various advanced statistics computed from box score data. Several studies, e.g., Teramoto and Cross (2010), Baghal (2012) or Page et al. (2007), explain the game outcomes using box scores and such advanced statistics, in particular the *four factors*. However, as this information in only known ex post, it is unclear whether these results can be exploited for forecasting purposes. A notable exception is the Markov model in Štrumbelj and Vračar (2012), in which the transition probabilities in a Markov chain model for basketball games are explained by the four factors.

The prediction of basketball games is the topic of Boulier and Stekler (1999), Caudill (2003) Loeffelhold et al. (2009), Rosenfeld et al. (2010), Stekler and Klein (2012), Štrumbelj and Vračar (2012), or Štrumbelj (2014). These predictions are done in very different settings and with quite different methodologies. In particular, forecasts are often based on team rankings, betting odds or statistical models. A common finding of many studies is that predictions based on betting markets are difficult to beat, thus implying efficiency of the betting markets; see also Steckler et al. (2010) and references therein on this issue.

This paper contributes to the aforementioned literature in several ways. Building on the benchmark linear model for team strengths, including parameters for the effect of the home court advantage and of playing back-to-back games, team specific volatility is introduced into the framework. The estimation and testing for heteroscedasticity is discussed. A second contribution is to consider models for time-varying team strengths. Two approaches are presented to this end. The first is allowing for structural breaks in the strength parameter of a specific team at an unknown point in time, whereas the second approach is a dynamic state space model in which the team strengths follow a Gaussian autoregressive process. The empirical analysis relies on a large dataset of eight NBA seasons. Estimates of teams strength and rankings, as well as the effect of the home court advantage and back-to-back games are compared across different models. Tests for heteroscedasticity are applied to the data providing some weak evidence against the assumption of equal error variances across teams. Furthermore, normality tests suggest that the residuals are normally distributed. Applying the time-varying models we find evidence for some structural breaks, but no evidence for persistent time-varying strength parameters. This is in line with the usual believe that the "hot hand" does not exist for teams; see the discussion in Camerer (1989) and Brown and Sauer (1993) on this issue. Finally, the forecasting performance of the proposed models is compared for a large number of regular season and playoff games. The model forecasts are compared to point spreads from the betting market and it turns out that this are a benchmark that is difficult to beat. The model based forecasts are also combined with the point spreads and the resulting forecast combinations often result in the best forecasts in the comparison.

The rest of the paper is structured as follow. In Section 2 the methodology is explained, Section 3 presents the empirical application and some conclusions are given in Section 4.

In the appendix estimation details for the dynamic state space model and additional estimation results can be found.

2 Methodology

Let y_{ijk} be the difference in scores of the home team i minus the away team j, where $k = 1, \ldots, n$ is the index of game k and n is the total number of games. The total number of teams is denoted by t and each team plays a total of K games, so that $n = t \times K$. A simple model for the outcome of the game is

$$y_{ijk} = \lambda + \alpha(B2B_i - B2B_j) + \beta_i - \beta_j + e_{ijk}, \tag{1}$$

where λ denotes the (constant) home advantage, $B2B_i$ is a dummy variable indicating whether team i plays back-to-back games, i.e., games on two consecutive days, with α the corresponding effect, and β_i and β_j denote the strength of teams i and j, respectively.[1] The error term e_{ijk} is assumed to be normally distributed with mean 0 and variance σ^2. Harville (2003) suggests accounting for the discreteness of the observed scores. However, normality tests below suggest that the residuals from model (1) and its extensions below are indeed normally distributed. Furthermore, normality of the error terms implies that the correction for blowout victories proposed in Harville (2003) is not necessary and would, in fact, lead to inefficient estimates given the fact that under normality ordinary least squares (OLS) is equivalent to the (asymptotically efficient) maximum likelihood estimator. We can state the model in matrix form letting $\boldsymbol{y}$ be the $n \times 1$ vector of spreads, $\boldsymbol{e}$ the $n \times 1$ vector of errors, $\boldsymbol{\beta} = [\lambda \ \alpha \ \beta_1 \ldots \beta_t]'$ the vector of coefficients and $\boldsymbol{X}$ the $n \times (t + 1)$ design matrix. A typical row of this matrix has 1 as its first element (for the home advantage), $B2B_i - B2B_j$ in the second column, 1 in column $i + 2$ and -1 in column $j + 2$ in the case that it corresponds to a game of team i (home) against team j (away). The remaining elements are equal to 0. Then the model is compactly given by

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{e}. \tag{2}$$

However, the matrix $\boldsymbol{X}$ is not of full rank, so for estimation one can remove the third column. This corresponds to the normalizing restriction $\beta_1 = 0$, meaning that the strength of the first team is set equal to zero. Without this restriction the parameter vector $\boldsymbol{\beta}$

[1] Here we made the assumption that the effect of playing back-to-back games is the same for the home and away team.

cannot be identified, as adding a constant to each team strength leads to an equivalent model. The parameters can be then estimated by OLS:

$$\hat{\boldsymbol{\beta}}_{OLS} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y}. \tag{3}$$

2.1 Heteroscedasticity

The model above assumes constant variance of the error term, i.e., $\boldsymbol{e} \sim N(\boldsymbol{0}, \sigma^2\boldsymbol{I})$, where $\boldsymbol{I}$ is the $n \times n$ identity matrix. Here we relax this assumption. Let the strength of team i in game k be given by

$$S_{ik} = \beta_i + e_{ik}, \tag{4}$$

where β_i is the constant component of the team strength and $e_{ik} \overset{iid}{\sim} N(0, \sigma_i^2)$ the team specific error term. Thus the strength of a team in a specific game consists of a constant component and an error term. A larger value of the error variance σ_i^2 implies that the corresponding team shows a more volatile performance. Then the outcome of the game is modeled as

$$y_{ijk} = \lambda + \alpha(B2B_i - B2B_j) + S_{ik} - S_{jk} = \lambda + \alpha(B2B_i - B2B_j) + \beta_i - \beta_j + \underbrace{e_{ik} - e_{jk}}_{e_{ijk}}. \tag{5}$$

Consequently, the baseline model (1) is obtained when $\sigma_i^2 = \sigma^2$ for all i. In matrix notation the model is the same as (2), but with $\mathrm{Cov}(\boldsymbol{e}) = \Omega \neq \sigma^2\boldsymbol{I}$. The matrix Ω is diagonal with typical element $\sigma_i^2 + \sigma_j^2$, corresponding to a game between team i and j.

The model can be estimated in two ways: Maximum likelihood estimation (MLE) or feasible generalized least squares (FGLS). MLE is straightforward since $e_{ijk} \sim N(0, \sigma_i^2 + \sigma_j^2)$ and the errors are independent. To estimate the model by FGLS first estimate (2) by OLS to obtain the residual vector $\hat{\boldsymbol{e}}$. Next, run the regression

$$\hat{\boldsymbol{e}}^2 = \boldsymbol{Z}\boldsymbol{\gamma} + \boldsymbol{\eta}, \tag{6}$$

where $\hat{\boldsymbol{e}}^2$ is the vector of squared residuals and the $n \times t$ matrix $\boldsymbol{Z}$ has typical row of zeros with entries of 1 in columns i and j if the observation corresponds to a game between teams i and j. The estimated parameter vector $\hat{\boldsymbol{\gamma}}$ in fact gives estimates for the team specific variances σ_i^2. The fitted values from (6), say $\hat{\sigma}_{ijk}^2$, make up the elements on the main diagonal of our estimate for the covariance matrix of the error terms $\hat{\Omega}$. Then the FGLS estimator is given by

$$\hat{\boldsymbol{\beta}}_{FGLS} = (\boldsymbol{X}'\hat{\Omega}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}'\hat{\Omega}^{-1}\boldsymbol{y}. \tag{7}$$

Consider testing the null hypothesis of homoscedasticity, i.e., a constant error variance accross teams,

$$H_0 : \sigma_i^2 = \sigma_j^2 \text{ for all } i \neq j. \tag{8}$$

There are two ways we can test this hypothesis. First, one could estimate model (5) by MLE and additionally estimate the model under the restriction of homoscedasticity. Let LL_0 be the log-likelihood under H_0 and LL_1 under the alternative. Then we can test H_0 using

$$LR = 2(LL_1 - LL0), \tag{9}$$

which follows a χ^2 distribution with $t-1$ degrees-of-freedom under the null. Alternatively, we can base our test on the regression (6). Let SSR_1 be the sum-of-squared residuals from this model and let SSR_0 be the residuals from regressing $\hat{\boldsymbol{e}}^2$ on a constant. Then we can test H_0 with the F-statistic

$$F = \frac{(SSR_0 - SSR_1)/(t-1)}{SSR_1/(n-t)}, \tag{10}$$

which is distributed $F(t-1, n-t)$.

In general, one may be interested in computing the probability that team i (the home team) wins a specific game. This can be computed as

$$\begin{aligned} P(\text{Team } i \text{ wins}) &= P(y_{ijk} > 0) = P(\lambda + \alpha B2B_i + S_{it} > \alpha B2B_j + S_{jk}) \\ &= P(\lambda + \alpha B2B_i + \beta_i + e_{ik} > \beta_j + \alpha B2B_j + e_{jk}) \\ &= P(e_{jk} - e_{ik} < \lambda + \alpha(B2B_i - B2B_j) + \beta_i - \beta_j) \\ &= P\left(\frac{e_{jk} - e_{ik}}{\sqrt{\sigma_i^2 + \sigma_j^2}} < \frac{\lambda + \alpha(B2B_i - B2B_j) + \beta_i - \beta_j}{\sqrt{\sigma_i^2 + \sigma_j^2}}\right) \\ &= \Phi\left(\frac{\lambda + \alpha(B2B_i - B2B_j) + \beta_i - \beta_j}{\sqrt{\sigma_i^2 + \sigma_j^2}}\right). \end{aligned} \tag{11}$$

2.2 Dynamic Modelling

Until now we have assumed that the strength parameter of a team is constant throughout the entire season. In this section we discuss two approaches to relax this assumption. In Section 2.2.1 we consider a model which permits a structural break for a certain team at an unknown point in time. Such a model may be extremely useful if one is interested in evaluating the impact of trade, injuries or changes in the coaching staff that occur with a season. Section 2.2.2 outlines a dynamic state space model, in which the strength of a team is a latent Gaussian autoregressive process. This model can be used to test the *hot hand* or *momentum* hypothesis that suggests that team strength varies over time and is persistent.

2.2.1 Structural change

Consider again the baseline model (1). Let the strength parameter be indexed by $k_i = 1, \ldots, 82$[2], i.e., we now have β_{i,k_i}. Consider testing the hypothesis

$$H_0 : \beta_{i,1} = \beta_{i,2} = \ldots = \beta_{i,82} \tag{12}$$

against the alternative

$$H_0 : \beta_{i,1} = \ldots = \beta_{i,k_i^*-1} \neq \beta_{i,k_i^*} = \ldots = \beta_{i,82}. \tag{13}$$

Thus we want to test constancy of the parameter β_{i,k_i} against the alternative of a single structural break at game k_i^*. If the time of the break k_i^* is known this can be done with the test of Chow (1960). This could be interesting if one is interested in testing whether the injury of a key player or a certain trade had an impact on the strength of a team. The test is based on the regression

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \delta D(k_i^*) + \boldsymbol{e}, \tag{14}$$

where $D(k_i^*)$ is the $i+1$th column of $\boldsymbol{X}$, i.e., the column corresponding to team i, with all entries for games 1 to k_i^* set to zero. Thus δ measures the change in team i's strength at the point in time corresponding to game k_i^*. Our null hypothesis is then equivalent to

$$H_0 : \delta = 0, \tag{15}$$

which can be tested with a standard t-test. Denote the t-statistic corresponding to $D(k_i^*)$ by $t_{k_i^*}$. Now consider the situation that the time of the structural break k_i^* is unknown. A test for the null hypothesis (12) with unknown breakpoint can be based on the statistic

$$F_{\sup} = \max_{k_i \in \Pi} t_{k_i}^2, \tag{16}$$

where Π is the set of potential breakpoints, which in our case excludes the first and last 10 games of the season. This truncation of the potential breakpoints is needed, because

[2]Note that each team plays 82 games per season, with the exception of lockout seasons such as the 2011-2012 season.

F_{sup} diverges at the boundary of the sample. Thus we test for a structural break with the maximum of the squared t-statistics[3] over all potential breakpoints. The asymptotic distribution of F_{sup} is non-standard and has been studied in Andrews (1993). However, instead of relying on asymptotic critical values we use a parametric bootstrap in our empirical application. This is achieved by repeatedly simulating from model (2) using the OLS estimates of $\boldsymbol{\beta}$ and σ^2 and drawing the error terms from a normal distribution. For each draw we estimate (14) and compute F_{sup}. Critical values are given by the empirical quantiles over the bootstrap distribution of F_{sup}.

Finally, a consistent estimate for the breakpoint is given by

$$k_i^* = \arg\max_{k_i \in \Pi} t_{k_i}^2 \tag{17}$$

and the change in team strength is the estimated value of δ from (14).

Note that the extension to multiple breakpoints is straightforward, as outlined in Bai and Perron (1998). However, given that each team plays only 82 regular season games we do not believe that one can identify more than one structural break in a given season. Nevertheless, if one in interested in jointly modeling multiple seasons this could be of interest.

2.2.2 A dynamic state space model

In the previous section we considered a model in which the (unconditional) strength of a team is allowed to shift in value at an unknown point in time. In this section we consider a model in which the strength of team i is a Gaussian autoregressive process of order one. The outcome of the game in this context is modeled by

$$y_{ijk} = \lambda + \alpha(B2B_i - B2B_j) + \beta_{i,k_i} - \beta_{j,k_j} + e_{ijk}, \tag{18}$$

where $e_{ijk} \overset{iid}{\sim} N(0, \sigma^2)$. The time-varying team strength evolves as

$$\beta_{i,k_i} = \mu_i + \phi_i \beta_{i,k_i-1} + \eta_{k_i}, \tag{19}$$

where $\eta_{k_i} \sim N(0, \sigma_{\eta_i}^2)$. Although this is a state space model and β_{i,k_i} is unobservable the estimation is relatively straightforward due to the fact that both e_k and η_{k_i} are normally distributed. The key difference to a standard state space model in time series analysis

[3]Recall that the squared t-statistic is equal to the F-statistic for testing a restriction on a single parameter.

is the fact that the observations are not equidistant in calendar time, and therefore the evolution of the strength is defined from game to game.[4] Nevertheless, the Kalman filter can be applied to estimate the model parameters and the strengths of the teams. The details on how this is done for this specific model are given in the appendix. We impose one set of restrictions to the model in order to reduce the number of free parameters, namely we restrict ϕ_i the same for all teams. Furthermore, we also consider imposing the restriction that $\sigma^2_{\eta_i}$ is the same for all teams, addressing the issue whether heteroscedasticity is still an issue when allowing for time-varying strength parameters. Again, a standard likelihood ratio test can be used to test this restriction.

3 Application

In this section we apply the models proposed in Section 2 to a large data set of NBA games covering the Seasons 2006-2007 until 2013-2014, thus a total of eight NBA seasons. The data was obtained from www.nbastuffer.com. Besides the outcomes of the games and betting odds[5], the data set contains further information that was not used in this study such as the box score, the starting lineups and some advanced basketball statistics.

In a typical regular season each of the 30 teams plays 82 games, resulting in a total of 1230 regular season games. An exception is the 2011-2012 lockout season in which each team played 66 games, implying a total of 990 regular season games. Furthermore, during the 2012-2013 season as a result of the bombing at the Boston marathon the game Boston vs. Indiana needed to be rescheduled and was eventually not played.

The rest of this section is structured as follows. In Section 3.1 we estimate models that assume a constant team strength within each season, apply the tests for heteroscedasticity, and compare the resulting rankings of the teams. In Section 3.2 the question of time-variation in team strength is addressed. Section 3.3 compares the forecasting performance of the models for both regular season and playoff games.

3.1 Static models and heteroscedasticity tests

In this section we address two questions. First, does the variance of the team strength differ between teams and, second, does the incorporation of heteroscedasticity influence the estimation of the team strength and the ranking of the teams.

[4]A model in which strength evolves in calendar time was also considered in a preliminary analysis.

[5]Based on www.scoresandodds.com.

Table 1: Tests for homescedasticity

	F-test	LR-test	LR-test dyn.
2006-2007	0.0449	0.1046	0.1194
2007-2008	0.1452	0.0257	0.0376
2008-2009	0.0098	0.0003	0.0004
2009-2010	0.0092	0.0022	0.0022
2010-2011	0.6020	0.4722	0.4534
2011-2012	0.1901	0.0437	0.0470
2012-2013	0.8204	0.3113	0.2956
2013-2014	0.0004	0.0000	0.0000

Note: Table 1 reports the p-values the null hypothesis of homescedastic errors against the alternative of team specific error variances based on the static model as described in Section 2.1, as well as on the dynamic state space model from Section 2.2.2.

Table 1 reports the p-values of the F-test and likelihood ratio test for the null hypothesis of homoscedasticity given in equations (9) and (10). The likelihood ratio test is additionally applied dynamic model characterized by equations (18) and (19). The tests were performed for each individual season. The results show that for most seasons the null hypothesis of homoscedasticity is rejected when looking at the tests individually. Only for two seasons, namely 2010-2011 and 2012-2013, the assumption of constant variance of the team strengths cannot be rejected. However, these results should be interpreted with certain care due to the fact that we are performing multiple hypothesis tests. The tests within each season are obviously highly correlated and we postulate that adjustments for multiple testing may be ignored. However, we are still testing for heteroscedasticity over eight seasons. Using a simple Bonferroni adjustment for each test individually suggests rejection when the p-value is below $0.05/8 = 0.00625$ when testing at $\alpha = 0.05$. This suggests rejection only in 3 out of 8 seasons.

Additionally, the Jarque-Bera (JB), Anderson-Darling (AD), Lilliefors, and Shapiro-Wilk (SW) normality tests were applied on the estimated residuals of the different models. In Table 2 we report the results for the model allowing for heteroscedasticity. In basically all cases normality cannot be rejected. Using the Bonferroni adjustment we cannot reject H_0 for any test. Given the large sample sizes these non-rejections are quite remarkable and confirm the normality assumption that is typically made in the literature.

The detailed estimation results can be found in Tables 3 and 4 concerning the estimated

Table 2: Normality tests on the residuals

	JB	Lilliefors	AD	SW
2006-2007	0.3636	0.4412	0.7242	0.6871
2007-2008	0.3395	>0.5	0.7152	0.5553
2008-2009	0.1775	>0.5	0.1578	0.1807
2009-2010	>0.5	>0.5	0.9306	0.9365
2010-2011	0.3927	>0.5	0.5056	0.688
2011-2012	>0.5	0.3968	0.5475	0.4182
2012-2013	>0.5	0.0100	0.0103	0.0985
2013-2014	>0.5	0.3421	0.4489	0.699

Note: Table 2 reports the p-values of the following tests for normality: Jarque-Bera (JB), Lillieofors, Anderson-Darling (AD) and Shapiro-Wilk (SW). The tests are applied to the residuals of the static heteroscedastic model (5) estimated by MLE.

effect of the home advantage and back-to-back games, as well as in Tables 8 to 15 in Appendix B for the estimated team strengths and variances. The parameter estimates show some differences between the different estimators and some slight differences in team rankings emerge when allowing for heteroscedasticty. The effect of the home advantage is estimated to be around 2.7 points per game, whereas the playing back-to-back games on average results in a disadvantage of about 1.8 points. Looking at the range of estimated team strengths it can be seen that the difference between the best and the worst team in the league implies an expected point difference between 13 and 20 points.

Looking at the variance estimates themselves no clear pattern emerges. High variances are possible both for successful and unsuccessful teams. Furthermore, no individual team is characterized by high or low volatility over several seasons.[6] However, factors that may explain the differences in volatility may be frequency and severity of injuries suffered by some teams, resting of older key players or changes in the coaching staff. We leave an investigation of this issue for future research.

[6]For example, one may expect a team such as the San Antonio Spurs that is known for its good management and that is consistently one of the top teams of the league to show a less variable performance than other organizations.

Table 3: Estimated home advantage

	OLS	GLS	MLE	Dynamic
2006-2007	2.65	2.52	2.50	2.49
2007-2008	3.15	3.12	3.11	3.11
2008-2009	2.99	3.23	3.24	3.21
2009-2010	2.14	2.16	2.16	2.16
2010-2011	2.87	2.80	2.81	2.82
2011-2012	2.59	2.74	2.54	2.52
2012-2013	2.96	3.03	3.03	3.05
2013-2014	2.29	2.22	2.29	2.21

Note: Table 3 reports the estimated effect of the home court advantage based on the models defined in equations (1), (5) and (18), denoted by OLS, GLS/MLE and Dynamic, respectively. GLS and MLE refer to the estimation method of the heteroscedastic model (5).

3.2 Dynamic modeling

The next step in the analysis is to consider the question whether team strength is varying over time within a given season. In Table 5 the estimated breakpoints in team strength using the approach outlined in Section 2.2.1 are reported. Several breakpoints are identified, although their number is relatively small considering that we are looking at a total of 30 teams over eight seasons. Note that only breakpoints significant at the 5% and 1% level are reported, as multiple hypothesis tests are performed, which has to be kept in mind when interpreting these results. However, the sample size of 82 regular season games per team is relatively small given the difficulty of the problem of endogenously identifying a change-point. Most of the estimated dates can be explained by specific events that took place around that particular date. For example, the breakpoint for the Miami Heat on Jan. 7, 2007 can be explained by an injury of their key player Shaquille O'Neal missing the first 30 games of the season. Another example is the change point on March 15, 2012 by Washington, which coincides exactly with a large three-team trade on that day.[7]

The next step in the analysis is the estimation of the dynamic state space model from Section 2.2.2. Intuitively this model seems a reasonable approach, as one would expect the strengths of teams to change throughout the course of a season due to injuries, trades, changes in coaching and team chemistry, etc. Surprisingly, the log-likelihood of the

[7]A list of the events associated with the estimated break dates can be provided by the author upon request.

Table 4: Estimated effect of back-to-back games

	OLS	GLS	MLE	Dynamic
2006-2007	-1.86	-1.80	-1.71	-1.75
2007-2008	-1.50	-1.20	-1.17	-1.27
2008-2009	-1.50	-1.28	-1.27	-1.32
2009-2010	-3.24	-3.27	-3.19	-3.18
2010-2011	-1.62	-1.91	-1.79	-1.83
2011-2012	-1.95	-1.76	-1.95	-1.98
2012-2013	-1.35	-1.40	-1.45	-1.44
2013-2014	-1.85	-1.70	-1.69	-1.73

Note: Table 4 reports the estimated effect of playing back-to-back games based on the models defined in equations (1), (5) and (18), denoted by OLS, GLS/MLE and Dynamic, respectively. GLS and MLE refer to the estimation method of the heteroscedastic model (5).

dynamic and static models are basically identical for all season and the point estimates for the persistence parameter ϕ is always quite close to 0. Furthermore, the smoothed and filtered estimates of the path of the team strengths looks rather erratic and do not suggest any persistence in team strength. In order to shed further light on the question of momentum in team strength we treat the residuals of the static model as panel data for each team over the course of the season and perform the Lagrange-multiplier test for autocorrelation by Baltagi and Li (1998). In all cases the null hypothesis of no-autocorrelation cannot be rejected[8]. Thus we can conclude that there is no evidence of persistent time-variation in the team strength within individuals seasons. Although this finding is surprising at first sight, it can be explained by the large degree of professionalism in the NBA and it is clear evidence against the believe in the *hot hand*.

3.3 Predictability

In this section I consider the problem of forecasting the game outcomes using the models described above. This is done in two settings. In Section 3.3.2 regular season games are predicted, whereas Section 3.3.2 focuses on playoff games. The forecasts are evaluated

[8]Detailed results for all unreported findings in this section are available from the author upon request.

Table 5: Estimated breakpoints in team strength

Season	Team	Date	Change	Record before	Record after
2006-2007	Golden State***	05.03.2007	11.3205	26-35	16-5
	Miami**	07.01.2007	8.1302	13-19	31-19
	Minnesota**	20.02.2007	-8.6933	25-27	7-23
2008-2009	Boston**	25.02.2009	-8.5731	46-12	16-8
	Portland**	16.03.2009	9.3325	41-25	13-3
2009-2010	Boston**	27.12.2009	-8.4914	23-5	27-27
	Indiana**	14.03.2010	9.757	21-44	11-6
2010-2011	Denver**	16.02.2011	7.8313	31-25	21-25
	Utah**	17.12.2010	-8.5803	18-8	21-35
2011-2012	Cleveland**	23.03.2012	-9.2674	17-24	4-18
	Minnesota**	28.03.2012	-11.7942	24-27	2-13
	New York**	12.03.2012	8.983	18-23	18-7
	Philadelphia**	18.01.2012	-12.2324	10-3	25-28
	Portland***	29.02.2012	-11.3512	18-16	10-22
	Washington***	15.03.2012	9.1682	9-32	11-14
2012-2013	Portland**	24.03.2013	-10.9225	33-36	0-13
	Sacramento**	26.02.2013	8.0524	19-38	9-16
	San Antonio***	08.03.2013	-11.122	48-14	10-10
	Washington***	07.01.2013	9.7524	4-28	25-25
2013-2014	Charlotte**	29.01.2014	7.5105	19-27	24-12
	Cleveland**	07.02.2014	8.1545	16-33	17-16
	Indiana***	22.01.2014	-11.3447	33-7	23-19

Note: Table 5 reports the estimated break date in team strength within each respective season (see Section 2.2.1 for the methodology) together with the estimated change in team strength, as well as the record before and after the change point. ** denotes statistical significance of the test statistic at the 5% level, and *** at the 1% level.

using three criteria. The first criterion is the mean square prediction error (MSE):

$$MSE = \sum_{k=1}^{n^*} (y_{ijk} - \hat{y}_{ijk})^2,$$

where n^* is the number of out-of-sample observations. The second criterion is the mean absolute prediction error (MAE),

$$MAE = \sum_{k=1}^{n^*} |y_{ijk} - \hat{y}_{ijk}|,$$

and the third criterion is the fraction of games in which the correct winner was predicted. Whereas the MSE is the obvious choice for the loss function given the fact that the error terms can safely be considered to be Gaussian, the other two criteria are easy to interpret.

The models considered in the forecasting exercise are the homoscedastic baseline model (OLS), the heteroscedastic model (Het.) estimated by MLE and the dynamic state space model (Dyn.). As a benchmark the Las Vegas opening spreads (Spr.) for bets on the games are considered. Furthermore, for all models we consider the combined forecasts of the models forecasts with the betting spreads. The forecasts are combined with equal weights, as a preliminary analysis suggested that the two types of forecasts have approximately the same variances and are highly correlated (> 0.9). Therefore more sophisticated weighting schemes do not appear to be sensible here; see Timmermann (2006) for extensions.

Besides comparing the predictions in terms of the aforementioned measures, additionally the model confidence set (MCS) by Hansen et al. (2011) is computed based on the MSE and MAE loss functions. The MCS is a set of models whose forecasting performance is not significantly different considering a certain loss function and it can be seen as an analogue to a confidence interval for competing (non-nested) models. Thus it acknowledges the fact that it is unlikely that a single model outperforms all the others, but that there are multiple models that perform equally well. The MCS is determined using a sequence of hypothesis tests. It eliminates inferior models based on the criterion of interest. P-values for the sequential tests are determined by bootstrap procedure as described in Hansen et al. (2011) and references therein. A size of 5% and 10000 bootstrap samples are used to compute the MCS.

3.3.1 Regular season

The forecasting performance for the regular season data is analyzed as follows. The first half of the regular season data, 615 games in a typical season, are used as the in-sample

Table 6: Forecast evaluation regular season

	OLS	Het.	Dyn.	Spr.	OLS-Spr.	Het.-Spr.	Dyn.-Spr.
2007							
MSE	152.19†	151.90†	152.08†	**141.89**	145.15	145.01	145.13
MAE	9.74†	9.74†	9.75†	**9.39**	9.49	9.48	9.49
Correct	0.644	0.646	0.647	0.662	0.662	**0.665**	0.662
2008							
MSE	136.71	137.48	136.80	135.12	133.27	133.69	**133.26**
MAE	9.28†	9.30†	9.29†	9.11	9.11	9.11	**9.11**
Correct	0.714	0.715	0.715	0.720	**0.722**	0.709	0.720
2009							
MSE	138.94†	138.44†	139.47†	135.40	135.33	**135.18**	135.38
MAE	9.18	9.15	9.19†	**9.03**	9.05	9.04	9.04
Correct	**0.712**	**0.712**	0.706	0.707	0.709	0.707	0.709
2010							
MSE	141.16	140.69	141.17	141.06	139.26	**138.99**	139.22
MAE	9.42	9.43	9.43	9.35	9.32	**9.32**	9.33
Correct	**0.698**	0.688	0.691	0.693	0.691	0.696	0.693
2011							
MSE	127.79	128.56	128.19†	126.98	**126.41**	126.70	126.62
MAE	8.91	8.96	8.92†	8.88	**8.86**	8.88	8.87
Correct	0.694	0.688	0.691	0.693	**0.699**	0.693	0.698
2012							
MSE	149.49†	148.15†	150.34†	**139.34**	141.95	141.25	142.17
MAE	9.57†	9.52†	9.56†	9.35	9.37	**9.33**	9.35
Correct	0.653	0.661	0.659	**0.681**	0.677	0.679	0.679
2013							
MSE	152.68†	153.82†	152.72†	**147.47**	148.42	148.96	148.35
MAE	9.60†	9.62†	9.60†	**9.38**	9.41	9.42	9.41
Correct	0.673	0.668	0.669	0.695	**0.697**	0.695	0.695
2014							
MSE	140.17†	140.09†	139.88†	**132.99**	134.48	134.32	134.35
MAE	9.22†	9.20†	9.21†	**8.94**	9.02	9.01	9.02
Correct	0.663	0.670	0.659	0.683	0.685	**0.686**	0.683
Pooled							
MSE	142.21†	142.25†	143.13†	**137.49**	137.93	137.93	138.37
MAE	9.36†	9.36†	9.38†	**9.17**	9.20	9.20	9.22
Correct	0.682	0.681	0.684	0.692	0.693	0.692	**0.694**

Note: Table 6 gives the predictive mean-square-error (MSE), mean-absolute-error (MAE) and fraction of correctly predicted outcomes for all games of the second half of each respective season based on recursively estimated model parameters. OLS refers to the homoscedastic model in (1), Het. to the heteroscedastic model in (5), Dyn. to the dynamic state space model in (18), and Spr. to the Las Vegas opening spreads. The remaining four columns refer to equally weighted forecast combinations. The results for the best performing model are presented in bold. A † implies that the corresponding model is not included in the 95% model confidence set.

period, whereas the remaining games constitute the out-of-sample period. The models are re-estimated using an expanding window scheme to produce forecasts for the full out-of-sample period. The results are presented in Table 6, where the results for the best performing model in each case are shown in bold. At the bottom of the table all the forecasts are pooled to give an overall picture of the forecasting performance. With respect to the MSE and MAE the betting spreads provide the best forecasts in most seasons and for the pooled forecasts. However, in several instances combined forecasts perform as well or better, and they are never much worse. The results from the model confidence set (excluded models marked by a †) show that the betting spreads and the combined forecasts are always included in the MCS, whereas the pure model based forecasts are often excluded. Concerning the fraction of correct predictions no single model stands out, but combined forecasts using either the homoscedastic or the heteroscedastic regression model can be recommended. Overall, between 66% and 72% of the game outcomes can be predicted correctly and it seems questionable that much better forecasts are possible, as a certain amount of randomness/unpredictability is an inherent part of sports.

3.3.2 Playoffs

For the forecast evaluation of the playoff games the complete regular season data is used as the training period, but the models are not re-estimated during playoff period. In the case of the dynamic state space, however, the information set is updated throughout the playoffs and the predicted values based on the Kalman filter are used as forecasts. Additionally to the models used for forecasting the regular season games we also consider the estimates considering the structural breaks from Table 5. Again, no single model dominates. The Las Vegas spreads provide good forecasts in many cases, in particular in terms of MSE and MAE. However, the regression based approaches and the forecast combinations outperform the spreads in several seasons. The model confidence set excludes only very few models and the excluded ones are always among the regression based approaches and once the combined forecast with the structural break model for the year 2013. In terms of predicting the correct outcomes different models perform well in each season. The differences in the percentage of correctly predicted games across the models can be up to 10% within one season. Overall, between 64% and almost 80% of all games are correctly predicted. In summary, for forecasting playoff games the potential to beat the betting spreads appears to be larger than for regular season games and relying on combined forecasts seems to be a sensible approach.

Table 7: Forecast evaluation playoffs

	OLS	Het.	SB	Dyn.	Spr.	OLS-Spr.	Het.-Spr.	SB-Spr.	Dyn.-Spr.
2007									
MSE	120.57	120.14	115.00	120.03	123.27	119.09	119.01	**114.17**	118.85
MAE	8.25	8.26	8.24	**8.23**	8.60	8.39	8.39	8.33	8.38
Correct	**0.760**	**0.760**	0.734	0.747	0.684	0.722	0.722	0.734	0.709
2008									
MSE	175.04	176.11	175.04	172.52	**163.88**	167.17	167.53	167.17	165.78
MAE	10.68	10.71	10.68	10.59	**10.13**	10.34	10.35	10.34	10.30
Correct	0.709	0.686	0.709	0.686	**0.733**	0.698	0.698	0.698	0.698
2009									
MSE	207.16	206.45	213.84	**206.19**	217.13	209.00	209.00	211.30	208.32
MAE	10.90	10.86	10.94	10.92	11.05	10.88	**10.85**	10.87	10.88
Correct	0.718	0.718	0.694	**0.729**	0.671	0.659	0.671	0.635	0.659
2010									
MSE	177.94	179.10	188.55	179.13	175.23	**174.53**	175.12	179.01	174.71
MAE	10.50	10.52	10.82	10.53	**10.16**	10.28	10.29	10.43	10.29
Correct	0.659	0.659	0.671	0.671	0.683	0.695	**0.707**	0.659	**0.707**
2011									
MSE	110.78	110.93	115.24	111.10	111.89	**110.44**	110.53	112.09	110.60
MAE	8.44	8.43	8.74	8.45	8.44	8.39	**8.39**	8.51	8.39
Correct	0.642	0.642	0.605	0.642	**0.654**	**0.654**	**0.654**	**0.654**	**0.654**
2012									
MSE	111.46†	111.35†	120.67†	111.70†	**97.96**	102.47	102.46	105.93	102.40
MAE	8.40†	8.40†	8.79†	8.44†	**7.71**	7.92	7.92	8.06	7.94
Correct	0.714	0.702	0.714	0.702	0.774	0.786	**0.798**	0.762	0.774
2013									
MSE	155.08	155.28	188.48†	153.61	**147.60**	150.07	150.23	164.80†	149.14
MAE	10.51	10.50	11.49†	10.46	**10.09**	10.29	10.29	10.76†	10.26
Correct	0.659	0.671	0.588	0.671	0.659	0.682	**0.694**	0.647	**0.694**
2014									
MSE	148.12	**146.17**	148.14	148.86	151.17	147.94	147.04	147.20	148.17
MAE	9.67	**9.61**	9.79	9.65	9.67	9.65	9.62	9.70	9.64
Correct	**0.640**	**0.640**	0.562	**0.640**	0.584	0.596	0.573	0.562	0.573
Pooled									
MSE	151.24†	151.14†	158.65†	150.85	148.96	148.046	148.06	150.70	**147.69**
MAE	9.69†	9.68†	9.96†	9.68	**9.50**	9.54	9.53	9.64	9.53
Correct	0.687	0.684	0.662	0.686	0.680	0.686	**0.689**	0.668	0.683

Note: Table 7 gives the predictive mean-square-error (MSE), mean-absolute-error (MAE) and fraction of correctly predicted outcomes for all games of the playoffs of each respective season based on model parameters estimated using the regular season data. OLS refers to the homoscedastic model in (1), Het. to the heteroscedastic model in (5), SB to the model allowing for a structural break, Dyn. to the dynamic state space model in (18), and Spr. to the Las Vegas opening spreads. The remaining four columns refer to equally weighted forecast combinations. The results for the best performing model are presented in bold. A † implies that the corresponding model is not included in the 95% model confidence set.

4 Conclusion

In this paper I have reconsidered the modeling of team strength in professional basketball. The standard model was extended by allowing for team specific error variances and time-variation in team strength. The latter was achieved by (i) testing for and dating structural breaks at unknown points during the season, and by (ii) allowing for autoregressive time varying latent team strengths. These models were applied to the NBA games in all eight seasons in the period 2006 until 2014. The results of the in-sample estimation suggest the presence of heteroscedasticity in most seasons and it is found that the rankings of the teams by their estimated strength can be different from the ones implied by the standard homoscedastic model. Additionally, normality of the residuals cannot be rejected, which favors the estimation of the models by least squares. This finding also implies that the modified least squares approach proposed in Harville (2003) that controls for blowout victories is not necessary, because the presence of such blowouts should result in outlying observations that would lead to rejection of normality tests. Furthermore, although there is some evidence of structural breaks in team strength that can typically be associated with specific events such as trades or injuries, no evidence for momentum is found when estimating a dynamic state space model for team strength. This is confirmed by the rejection of tests for no autocorrelation on the residuals of the static models. Thus this paper provides further evidence against the presence of *momentum* or *hot hand* effects.

Besides the methods presented in this paper several other models were considered that were not able to improve the model fit. In particular, a model treating offensive and defensive strength separately in both a static and dynamic setting did not yield a better fit than its counterpart considering only one strength parameter. Furthermore, instead of the dynamic state space model, an autoregressive observation driven approach for team strength in which the residuals of the previous game were allowed to drive the current team strength was considered. Due to the absence of any evidence for the *hot hand* belief it is not surprising that such a model could not outperform simpler static models.

The forecasting performance of the models was evaluated using regular season and playoff games over all eight seasons. These finding confirm the common theme in the literature on sports forecasting: it is difficult to beat the betting markets, which indicates that they efficient. However, combining the model based forecasts with betting spreads sometimes leads to better forecasts and the model confidence sets imply that the combined forecasts are statically not worse than the one based solely on betting spreads.

Future research should address the question whether advanced basketball statistics suggested in Kubatko et al. (2007) can be used to improve model based forecasts and whether these statistics themselves are predictable. Furthermore, more detailed information concerning injuries or suspensions of key players can be incorporated into the models for forecasting purposes. Finally, it could be interesting to search for factors that can explain the varying variances of each team's strength parameter.

A Implementation of the Kalman filter

Let $\beta_{i,k_i|k_i-1}$ be the predicted team strength of team i for game k_i conditional on the information at game $k_i - 1$, whereas $\beta_{i,k_i|k_i}$ denotes the updated strength conditional on information up to game k_i. The variance of β_{i,k_i} conditional on information at game k_i-1 is denoted as $P_{i,k_i|k_i-1}$, whereas the updated variance of team i is $P_{i,k_i|k_i}$. Then the steps of the Kalman filter for game k between teams i and j with outcome y_{ijk}, being games k_i and k_j for the teams, respectively, are as follows.

Prediction step:

$$\begin{aligned}
\beta_{i,k_i|k_i-1} &= \mu_i + \phi\beta_{i,k_i-1|k_i-1} \\
\beta_{i,k_j|k_j-1} &= \mu_j + \phi\beta_{i,k_j-1|k_j-1} \\
P_{i,k_i|k_i-1} &= \phi^2 P_{i,k_i-1|k_i-1} + \sigma^2_{\eta_i} \\
P_{i,k_j|k_j-1} &= \phi^2 P_{i,k_j-1|k_j-1} + \sigma^2_{\eta_j}
\end{aligned}$$

Observation step:

$$\begin{aligned}
\hat{y}_{ijk} &= \lambda + \alpha(B2B_i - B2B_j) + \beta_{i,k_i|k_i-1} - \beta_{i,k_j|k_j-1} \\
V_{ijk} &= P_{i,k_i|k_i-1} + P_{i,k_j|k_j-1} + \sigma^2 \\
\hat{e}_{ijk} &= y_{ijk} - \hat{y}_{ijk}
\end{aligned}$$

Updating step:

$$\begin{aligned}
\beta_{i,k_i|k_i} &= \beta_{i,k_i|k_i-1} + \hat{e}_{ijk} P_{i,k_i|k_i-1}/V_{ijk} \\
\beta_{i,k_j|k_j} &= \beta_{i,k_j|k_j-1} - \hat{e}_{ijk} P_{i,k_j|k_j-1}/V_{ijk} \\
P_{i,k_i|k_i} &= P_{i,k_i|k_i-1} - P^2_{i,k_i|k_i-1}/V_{ijk} \\
P_{i,k_j|k_j} &= P_{i,k_j|k_j-1} - P^2_{i,k_j|k_j-1}/V_{ijk}
\end{aligned}$$

The initial values are set to $\beta_{i,1|0} = \mu_i/(1-\phi)$ and $P_{i,1|0} = \sigma^2_{\eta_i}/(1-\phi^2)$. The log-likelihood contribution of the kth game is given by

$$\ln L_k = \frac{1}{2}\ln(2\pi) + \frac{1}{2}\ln(V_{ijk}) + \frac{\hat{e}^2_{ijk}}{2V_{ijk}}$$

Finally, if one is interested in the estimates of the strength conditional on the information of the whole sample the Kalman smoother should be applied. Smoothed state estimates, denoted as $\beta_{i,k_i|K}$, are obtained by iterating the following recursion on the whole sample going from the last to the first game:

$$\beta_{i,k_i|K} = \beta_{i,k_i|k_i} + \phi \frac{P_{i,k_i|k_i}}{P_{i,k_i+1|k_i}} (\beta_{i,k_i+1|K} - \beta_{i,k_i+1|k_i}).$$

B Estimated team strenghts, rankings and error variances

Table 8: Ranking, strength and team specific variances 2006-2007

2006-2007	rank OLS	rank GLS	rank MLE	$\hat{\beta}_{OLS}$	$\hat{\beta}_{FGLS}$	$\hat{\beta}_{MLE}$	$\hat{\sigma}^2_{GLS}$	$\hat{\sigma}^2_{MLE}$
San Antonio	1	1	1	13.10	13.01	13.00	108.02	101.59
Phoenix	2	3	3	12.11	12.07	12.16	56.05	54.06
Dallas	3	2	2	11.99	12.37	12.37	72.68	56.22
Houston	4	4	4	9.76	9.53	9.64	99.84	105.32
Chicago	5	5	5	9.32	9.24	9.30	115.49	113.83
Detroit	6	6	6	8.54	9.06	9.10	60.80	61.67
Cleveland	7	7	7	8.17	8.17	8.14	38.67	41.28
Utah	8	8	8	7.79	7.80	7.85	76.36	71.81
Denver	9	9	9	6.49	6.20	6.24	58.22	68.04
Toronto	10	10	10	5.53	5.35	5.42	38.14	44.81
LA Lakers	11	12	13	5.15	4.85	4.83	68.34	71.86
Orlando	12	13	12	5.07	4.79	4.97	78.85	77.58
Golden State	13	11	11	4.80	4.97	5.06	140.24	135.96
LA Clippers	14	14	14	4.77	4.63	4.70	72.64	73.78
Washington	15	15	15	3.99	4.29	4.19	64.38	68.32
New Jersey	16	16	16	3.71	4.15	4.11	35.10	30.97
New Orleans	17	17	17	3.62	3.85	3.86	34.42	29.90
Sacramento	18	19	19	3.62	3.36	3.31	42.10	44.29
Miami	19	18	18	3.48	3.39	3.42	108.73	95.24
Seattle	20	20	20	2.39	2.77	2.70	61.07	75.66
Indiana	21	21	21	2.37	1.96	2.06	49.54	62.32
New York	22	24	24	1.64	1.49	1.48	48.38	44.66
Minnesota	23	23	23	1.64	1.50	1.53	49.50	50.04
Philadelphia	24	22	22	1.56	1.77	1.85	62.45	59.34
Boston	25	25	25	1.00	1.11	1.12	49.73	51.67
Portland	26	26	26	0.88	0.92	1.00	73.72	76.82
Charlotte	27	27	27	0.74	0.81	0.89	95.03	84.51
Memphis	28	29	29	0.35	0.09	0.17	31.10	34.37
Milwaukee	29	28	28	0.32	0.16	0.24	71.03	63.65
Atlanta	30	30	30	0.00	0.00	0.00	61.93	72.82

Note: Table 8 presents the estimated ranking, team strengths and team specific error variances based on models (1) and (5). The heteroscedastic model is estimated either by FGLS or by MLE.

Table 9: Ranking, strength and team specific variances 2007-2008

2007-2008	rank OLS	rank GLS	rank MLE	$\hat{\beta}_{OLS}$	$\hat{\beta}_{FGLS}$	$\hat{\beta}_{MLE}$	$\hat{\sigma}^2_{GLS}$	$\hat{\sigma}^2_{MLE}$
Boston	1	1	1	11.50	11.90	11.87	55.81	51.78
LA Lakers	2	2	2	9.56	9.58	9.35	79.83	80.87
Utah	3	4	4	9.00	8.74	8.46	103.24	110.39
Detroit	4	3	3	8.92	9.29	9.28	93.34	103.61
New Orleans	5	5	5	7.61	8.17	8.29	99.04	103.40
San Antonio	6	6	6	7.35	7.97	8.23	40.64	33.34
Phoenix	7	7	7	7.16	7.86	7.93	53.19	59.11
Houston	8	8	8	6.91	7.35	7.68	40.05	28.21
Orlando	9	10	10	6.90	6.96	7.05	62.59	62.22
Dallas	10	9	9	6.86	7.34	7.42	38.09	52.75
Denver	11	11	11	5.84	6.19	6.23	95.84	106.00
Toronto	12	12	12	4.61	5.16	5.24	106.56	110.29
Golden State	13	13	13	4.58	4.93	5.09	80.11	98.97
Philadelphia	14	14	14	2.22	2.58	2.85	79.68	103.89
Cleveland	15	15	15	1.85	2.30	2.37	41.92	23.51
Portland	16	17	17	1.76	1.49	1.00	27.50	4.43
Washington	17	16	16	1.45	1.97	1.99	103.12	112.76
Sacramento	18	18	18	0.41	0.73	0.97	56.56	60.59
Indiana	19	19	19	0.32	0.43	0.47	41.09	49.84
Atlanta	20	20	20	0.00	0.00	0.00	25.03	26.19
Chicago	21	21	21	-1.01	-0.78	-0.69	70.84	67.22
Charlotte	22	22	22	-2.25	-1.72	-1.43	59.92	57.45
New Jersey	23	23	23	-3.04	-2.41	-2.31	56.29	80.30
Memphis	24	24	24	-3.66	-3.46	-3.43	54.23	67.49
Minnesota	25	25	25	-3.97	-3.70	-3.85	60.24	50.14
New York	26	28	27	-4.36	-4.38	-4.17	74.28	61.13
LA Clippers	27	26	28	-4.45	-3.99	-4.18	50.37	53.96
Milwaukee	28	27	26	-4.89	-4.31	-3.86	58.84	37.19
Seattle	29	29	29	-5.97	-5.30	-4.91	66.05	61.89
Miami	30	30	30	-6.39	-5.37	-5.11	51.55	40.90

Note: Table 9 presents the estimated ranking, team strengths and team specific error variances based on models (1) and (5). The heteroscedastic model is estimated either by FGLS or by MLE.

Table 10: Ranking, strength and team specific variances 2008-2009

2008-2009	rank OLS	rank GLS	rank MLE	$\hat{\beta}_{OLS}$	$\hat{\beta}_{FGLS}$	$\hat{\beta}_{MLE}$	$\hat{\sigma}^2_{GLS}$	$\hat{\sigma}^2_{MLE}$
Cleveland	1	1	1	7.00	6.39	6.32	49.44	53.19
Boston	2	3	3	5.71	5.16	5.20	65.90	66.93
LA Lakers	3	2	2	5.55	5.32	5.29	28.65	43.18
Orlando	4	4	4	4.66	4.67	4.51	87.02	97.93
Portland	5	5	5	3.08	2.89	2.88	77.53	90.15
Houston	6	6	7	1.94	1.41	1.23	23.53	20.29
San Antonio	7	8	6	1.60	1.31	1.27	42.19	59.02
Denver	8	7	8	1.40	1.37	1.20	79.61	64.37
Utah	9	9	9	0.57	0.71	0.69	42.54	45.46
Atlanta	10	11	10	0.00	0.00	0.00	33.54	27.66
Dallas	11	10	11	-0.02	0.00	-0.15	85.29	87.54
Phoenix	12	12	12	-0.14	-0.70	-0.68	78.29	71.61
New Orleans	13	13	13	-0.42	-0.74	-0.93	45.33	50.76
Miami	14	14	14	-1.19	-1.10	-1.18	30.28	35.36
Philadelphia	15	15	15	-1.52	-1.45	-1.45	41.77	40.06
Chicago	16	16	16	-1.83	-1.97	-1.85	61.04	47.01
Detroit	17	17	18	-1.99	-2.16	-2.42	36.99	25.23
Indiana	18	18	17	-2.47	-2.25	-2.36	21.82	18.98
Milwaukee	19	19	19	-2.70	-2.92	-3.13	71.48	61.47
Charlotte	20	20	20	-2.90	-3.15	-3.27	43.59	62.13
New York	21	21	21	-4.02	-4.13	-4.25	94.84	89.81
New Jersey	22	22	22	-4.08	-4.31	-4.35	142.99	144.70
Toronto	23	23	23	-4.17	-4.42	-4.49	70.52	83.92
Golden State	24	24	24	-5.59	-5.84	-6.00	62.30	57.08
Minnesota	25	25	25	-6.41	-6.26	-6.26	74.48	79.85
Memphis	26	26	26	-6.84	-6.71	-6.57	59.49	48.74
Oklahoma City	27	27	27	-7.71	-7.87	-7.87	69.34	56.48
Washington	28	28	28	-8.75	-8.97	-8.97	52.46	53.02
Sacramento	29	29	29	-10.16	-10.14	-10.29	91.82	82.26
LA Clippers	30	30	30	-10.27	-10.35	-10.33	141.39	150.31

Note: Table 10 presents the estimated ranking, team strengths and team specific error variances based on models (1) and (5). The heteroscedastic model is estimated either by FGLS or by MLE.

Table 11: Ranking, strength and team specific variances 2009-2010

2009-2010	rank OLS	rank GLS	rank MLE	$\hat{\beta}_{OLS}$	$\hat{\beta}_{FGLS}$	$\hat{\beta}_{MLE}$	$\hat{\sigma}^2_{GLS}$	$\hat{\sigma}^2_{MLE}$
Orlando	1	1	1	2.76	2.45	2.58	43.39	60.18
Cleveland	2	2	2	1.70	1.27	1.34	11.14	16.07
LA Lakers	3	5	4	0.86	0.29	0.32	61.10	65.79
Utah	4	3	3	0.84	0.56	0.50	64.69	65.95
San Antonio	5	4	5	0.62	0.31	0.14	48.89	33.43
Phoenix	6	7	7	0.22	-0.21	-0.26	70.21	64.12
Atlanta	7	6	6	0.00	0.00	0.00	50.85	48.77
Denver	8	8	8	-0.23	-0.54	-0.66	65.71	73.76
Oklahoma City	9	9	10	-0.63	-1.04	-1.18	44.68	42.03
Boston	10	10	11	-0.88	-1.19	-1.20	81.11	95.02
Portland	11	11	9	-0.97	-1.27	-1.08	43.61	42.96
Dallas	12	12	12	-1.70	-2.22	-2.20	134.36	106.67
Miami	13	13	13	-2.48	-2.85	-2.77	128.09	137.77
Milwaukee	14	15	15	-3.21	-3.53	-3.65	52.15	56.74
Charlotte	15	14	14	-3.37	-3.40	-3.35	65.44	64.72
Houston	16	16	16	-4.20	-4.33	-4.37	77.32	72.54
Memphis	17	18	18	-5.65	-5.96	-5.95	67.38	61.09
Chicago	18	17	17	-5.85	-5.93	-5.75	89.01	86.13
Toronto	19	19	19	-6.07	-6.43	-6.45	74.90	82.45
New Orleans	20	21	21	-6.73	-7.73	-7.75	25.26	30.10
Indiana	21	20	20	-7.24	-6.89	-6.82	78.87	65.10
Golden State	22	22	22	-7.50	-7.77	-7.81	83.35	91.18
Philadelphia	23	23	23	-8.36	-8.29	-8.13	55.62	53.52
Sacramento	24	25	25	-8.43	-8.70	-8.73	23.13	16.91
New York	25	24	24	-8.48	-8.61	-8.56	134.00	124.41
Detroit	26	26	26	-9.27	-8.81	-8.82	61.91	56.23
Washington	27	27	27	-9.27	-9.03	-9.27	36.37	42.64
LA Clippers	28	28	28	-10.35	-10.86	-10.92	91.72	101.09
New Jersey	29	29	29	-13.18	-13.10	-13.14	35.84	46.06
Minnesota	30	30	30	-13.52	-13.88	-13.91	62.89	61.86

Note: Table 11 presents the estimated ranking, team strengths and team specific error variances based on models (1) and (5). The heteroscedastic model is estimated either by FGLS or by MLE.

Table 12: Ranking, strength and team specific variances 2010-2011

2010-2011	rank OLS	rank GLS	rank MLE	$\hat{\beta}_{OLS}$	$\hat{\beta}_{FGLS}$	$\hat{\beta}_{MLE}$	$\hat{\sigma}^2_{GLS}$	$\hat{\sigma}^2_{MLE}$
Miami	1	1	1	7.81	7.86	7.87	75.27	79.33
Chicago	2	2	2	7.49	7.16	7.18	58.24	42.75
LA Lakers	3	3	3	7.11	6.97	7.05	90.79	92.06
San Antonio	4	4	4	6.81	6.91	6.98	43.77	35.18
Orlando	5	5	5	5.96	6.36	6.35	66.72	72.15
Boston	6	8	7	5.86	5.47	5.56	42.20	53.41
Denver	7	6	6	5.77	5.60	5.68	72.69	79.27
Dallas	8	7	8	5.42	5.50	5.50	34.63	36.48
Oklahoma City	9	9	9	4.68	4.44	4.55	38.93	44.96
Memphis	10	11	10	3.59	3.50	3.56	54.61	56.49
Houston	11	10	11	3.40	3.52	3.45	27.41	27.83
Portland	12	12	12	2.89	3.13	3.19	62.64	64.19
New Orleans	13	14	14	2.21	2.17	2.24	76.22	71.19
Philadelphia	14	13	13	2.16	2.32	2.27	74.79	77.40
New York	15	15	15	1.66	1.80	1.78	61.53	68.96
Phoenix	16	16	16	0.42	0.44	0.52	45.23	41.52
Atlanta	17	18	17	0.00	0.00	0.00	113.23	106.98
Milwaukee	18	17	18	-0.05	0.01	-0.07	65.30	62.14
Indiana	19	19	19	-0.30	-0.11	-0.18	75.16	73.22
Utah	20	20	20	-0.47	-0.59	-0.57	71.07	63.25
Golden State	21	21	21	-0.92	-1.07	-1.10	57.41	56.26
LA Clippers	22	22	22	-1.59	-1.72	-1.75	37.73	44.60
Detroit	23	23	23	-2.75	-2.73	-2.68	31.86	33.36
Charlotte	24	24	24	-3.15	-3.14	-3.06	52.33	51.34
Sacramento	25	25	25	-3.83	-4.11	-4.12	62.82	59.34
Minnesota	26	26	26	-5.04	-5.00	-4.92	67.89	69.54
Toronto	27	28	28	-5.22	-5.12	-5.11	56.07	61.79
New Jersey	28	27	27	-5.22	-5.04	-5.03	25.42	29.89
Washington	29	29	29	-6.34	-6.34	-6.30	65.92	68.60
Cleveland	30	30	30	-7.97	-7.74	-7.70	65.37	51.23

Note: Table 12 presents the estimated ranking, team strengths and team specific error variances based on models (1) and (5). The heteroscedastic model is estimated either by FGLS or by MLE.

Table 13: Ranking, strength and team specific variances 2011-2012

2011-2012	rank OLS	rank GLS	rank MLE	$\hat{\beta}_{OLS}$	$\hat{\beta}_{FGLS}$	$\hat{\beta}_{MLE}$	$\hat{\sigma}^2_{GLS}$	$\hat{\sigma}^2_{MLE}$
Chicago	1	1	1	4.87	4.86	4.70	80.76	102.14
San Antonio	2	2	2	4.68	4.57	4.59	80.67	66.34
Oklahoma City	3	3	3	4.04	3.90	3.83	15.03	8.47
Miami	4	4	4	3.44	2.95	3.00	83.12	86.04
Philadelphia	5	5	5	1.04	0.75	0.66	113.90	106.65
Denver	6	7	8	0.54	0.13	0.05	101.49	91.00
LA Clippers	7	8	7	0.20	0.09	0.11	44.66	68.77
Indiana	8	6	6	0.10	0.16	0.12	51.16	56.59
Atlanta	9	9	9	0.00	0.00	0.00	59.30	74.76
New York	10	13	14	-0.14	-0.76	-0.98	77.50	71.31
Memphis	11	10	10	-0.14	-0.46	-0.29	32.94	26.90
Boston	12	11	11	-0.18	-0.53	-0.69	68.54	66.14
LA Lakers	13	12	12	-0.46	-0.65	-0.70	21.65	32.02
Dallas	14	14	13	-0.64	-0.91	-0.96	72.24	69.97
Utah	15	16	16	-1.58	-2.14	-2.05	29.92	59.72
Houston	16	18	18	-2.03	-2.64	-2.51	38.58	40.34
Orlando	17	17	17	-2.04	-2.43	-2.40	74.61	86.46
Phoenix	18	15	15	-2.14	-1.97	-2.01	69.18	65.41
Milwaukee	19	20	20	-2.76	-3.60	-3.56	47.52	49.75
Portland	20	19	19	-2.76	-3.08	-3.05	140.57	136.12
Minnesota	21	21	21	-4.37	-4.76	-4.75	62.14	49.21
Golden State	22	23	23	-5.28	-5.76	-6.00	83.14	74.56
New Orleans	23	22	22	-5.68	-5.36	-5.39	26.12	14.38
Toronto	24	24	24	-6.17	-6.13	-6.13	89.52	97.46
Sacramento	25	25	25	-7.38	-7.05	-7.24	81.56	95.08
Detroit	26	26	27	-7.69	-8.01	-8.02	75.94	67.74
Washington	27	27	26	-7.74	-8.02	-7.93	98.56	99.42
New Jersey	28	28	28	-8.76	-9.10	-9.03	54.44	30.23
Cleveland	29	29	29	-9.56	-9.32	-9.20	83.20	94.20
Charlotte	30	30	30	-16.64	-16.74	-16.64	68.05	56.44

Note: Table 13 presents the estimated ranking, team strengths and team specific error variances based on models (1) and (5). The heteroscedastic model is estimated either by FGLS or by MLE.

Table 14: Ranking, strength and team specific variances 2012-2013

2012-2013	rank OLS	rank GLS	rank MLE	$\hat{\beta}_{OLS}$	$\hat{\beta}_{FGLS}$	$\hat{\beta}_{MLE}$	$\hat{\sigma}^2_{GLS}$	$\hat{\sigma}^2_{MLE}$
Oklahoma City	1	1	1	9.26	8.79	8.76	58.51	66.72
Miami	2	2	2	7.14	6.92	6.87	46.89	46.97
San Antonio	3	3	4	6.87	6.56	6.51	92.12	86.43
LA Clippers	4	4	3	6.63	6.52	6.54	87.72	95.09
Denver	5	5	5	5.46	5.04	5.09	39.67	30.69
Memphis	6	6	6	4.43	4.27	4.38	34.78	29.92
New York	7	7	7	3.90	4.06	4.06	82.61	90.31
Houston	8	8	8	3.88	3.60	3.69	99.96	114.10
Indiana	9	9	9	3.41	3.18	3.19	58.66	59.77
LA Lakers	10	10	10	1.67	1.46	1.52	38.99	43.93
Brooklyn	11	11	12	1.50	1.12	1.13	81.36	85.01
Golden State	12	12	11	1.42	1.12	1.21	68.83	51.55
Utah	13	13	14	0.39	0.26	0.22	41.42	38.46
Chicago	14	14	13	0.06	0.13	0.26	98.39	106.48
Atlanta	15	15	15	0.00	0.00	0.00	67.89	60.40
Dallas	16	16	16	-0.21	-0.43	-0.28	72.96	76.30
Boston	17	17	17	-0.39	-0.44	-0.39	64.77	53.86
Minnesota	18	19	19	-1.64	-2.16	-2.17	57.32	59.70
Milwaukee	19	18	18	-1.73	-2.00	-1.88	44.48	45.40
Toronto	20	20	20	-1.80	-2.24	-2.27	78.95	81.10
Portland	21	21	21	-2.47	-2.72	-2.64	72.96	57.05
Washington	22	22	22	-2.68	-2.74	-2.72	37.31	41.78
New Orleans	23	24	24	-2.96	-3.23	-3.01	66.26	79.53
Philadelphia	24	23	23	-3.32	-3.07	-2.88	50.19	52.97
Sacramento	25	27	27	-4.21	-4.54	-4.48	92.03	88.06
Detroit	26	25	25	-4.25	-4.13	-3.99	91.64	97.29
Cleveland	27	26	26	-4.78	-4.47	-4.28	40.23	28.42
Phoenix	28	28	28	-5.67	-5.73	-5.55	77.79	76.98
Orlando	29	29	29	-7.05	-7.04	-7.01	92.11	102.29
Charlotte	30	30	30	-9.19	-9.24	-9.27	56.55	58.35

Note: Table 14 presents the estimated ranking, team strengths and team specific error variances based on models (1) and (5). The heteroscedastic model is estimated either by FGLS or by MLE.

Table 15: Ranking, strength and team specific variances 2013-2014

2013-2014	rank OLS	rank GLS	rank MLE	$\hat{\beta}_{OLS}$	$\hat{\beta}_{FGLS}$	$\hat{\beta}_{MLE}$	$\hat{\sigma}^2_{GLS}$	$\hat{\sigma}^2_{MLE}$
San Antonio	1	1	1	8.84	9.64	9.66	79.24	70.32
LA Clippers	2	2	2	7.92	8.21	8.27	82.13	74.79
Oklahoma City	3	3	3	7.48	7.16	7.36	59.27	69.95
Houston	4	5	5	5.86	6.16	6.37	70.48	54.04
Golden State	5	4	4	5.82	6.39	6.50	56.22	55.69
Portland	6	7	7	5.20	5.14	5.25	49.87	50.38
Miami	7	6	6	5.11	6.10	6.20	64.17	68.75
Indiana	8	9	9	4.53	4.65	4.78	88.06	85.47
Phoenix	9	8	8	3.95	4.71	5.00	43.55	36.20
Minnesota	10	12	11	3.94	4.17	4.16	75.88	63.61
Dallas	11	10	10	3.68	4.34	4.27	22.79	19.12
Toronto	12	11	12	3.28	4.26	4.14	0.77	2.45
Memphis	13	13	13	2.93	2.52	2.84	43.24	65.23
Chicago	14	14	14	1.81	2.07	2.21	95.17	111.65
Washington	15	15	15	1.47	1.60	1.78	39.32	50.19
Charlotte	16	16	16	0.04	0.82	0.84	67.37	84.83
Atlanta	17	19	20	0.00	0.00	0.00	46.42	59.23
New York	18	20	19	-0.46	-0.05	0.10	154.23	157.26
Denver	19	17	17	-0.58	0.30	0.45	98.23	98.42
Brooklyn	20	18	18	-0.63	0.17	0.22	109.36	89.03
Sacramento	21	21	21	-1.07	-0.33	-0.10	76.78	73.75
New Orleans	22	22	22	-1.20	-1.11	-1.30	23.80	24.21
Cleveland	23	23	23	-2.94	-2.54	-2.43	85.04	74.78
Detroit	24	24	24	-3.28	-3.17	-3.05	71.65	64.93
Boston	25	25	25	-4.05	-3.25	-3.09	46.00	36.98
LA Lakers	26	26	26	-4.31	-4.08	-3.99	102.43	101.73
Orlando	27	27	27	-4.99	-4.65	-4.67	32.04	24.60
Utah	28	28	28	-5.36	-5.19	-5.05	77.77	85.48
Milwaukee	29	29	29	-7.51	-7.02	-6.98	10.07	15.82
Philadelphia	30	30	30	-9.91	-9.72	-9.57	96.27	101.89

Note: Table 15 presents the estimated ranking, team strengths and team specific error variances based on models (1) and (5). The heteroscedastic model is estimated either by FGLS or by MLE.

References

Andrews, D. W. K. (1993). Tests for parameter instability and structural change with unknown change point. *Econometrica 61*(4), 821–856.

Baghal, T. (2012). Are the "four factors" indicators of one factor? An application of structural equation modeling methodology to nba data in prediction of winning percentage. *Journal of Quantitative Analysis in Sports 8*(1), Article: 4.

Bai, J. and P. Perron (1998). Estimating and testing linear models with multiple structural changes. *Econometrica 66*(1), 47–78.

Baltagi, B. H. and Q. Li (1998). Testing AR(1) against MA(1) disturbances in an error component model. *Econometrica 66*(1), 47–78.

Boulier, B. L. and H. O. Stekler (1999). Are sports seedings good predictors?: An evaluation. *International Journal of Forecasting 15*, 83–91.

Bradley, R. A. and M. E. Terry (1952). The rank analysis of incomplete designs, 1. The method of paired comparisons. *Biometrika 39*, 324–345.

Brown, W. O. and R. D. Sauer (1993). Does the basketball market believe in the hot hand? Comment. *American Economic Review 83*, 1377–1386.

Camerer, C. F. (1989). Does the basketball market believe in the hot hand? *American Economic Review 79*, 1257–1261.

Carlin, B. P. (1996). Improved NCAA basketball tournament modeling via point spread and team strength information. *The American Statistician 50*, 39–43.

Caudill, S. B. (2003). Predicting discrete outcomes with the maximum score estimator: The case of the NCAA men's basketball tournament. *International Journal of Forecasting 19*, 313–317.

Chow, G. (1960). Tests of equality between sets of coefficients in two linear regressions. *Econometrica 28*, 591–605.

David, H. A. (1959). Tournaments and paired comparisons. *Biometrika 46*, 139–149.

Elo, A. E. (1978). *The rating of chess players past and present.* New York: Arco.

Entine, O. A. and D. S. Small (2008). The role of rest in the nba home-court advantage. *Journal of Quantitative Analysis in Sports 4*(2), Article: 6.

Fahrmeir, L. and G. Tutz (1994). Dynamic stochastic models for time-dependent ordered paired comparison systems. *Journal of the American Statistical Association 89*, 1438–1449.

Glickman, M. E. (1993). *Paired comparison models with time-varying parameters.* Phd dissertation, Department of Statistics, Harvard University, Cambridge.

Glickman, M. E. (1999). Parameter estimation in large dynamic paired comparison experiments. *Applied Statistics 48*, 377–394.

Glickman, M. E. (2001). Dynymic paired comparison models with stochastic variances. *Journal of Applied Statistics 28*, 673–689.

Glickman, M. E. and H. S. Stern (1998). A state-space model for national football league scores. *Journal of the American Statistical Association 93*, 25–35.

Hansen, P. R., A. Lunde, and J. M. Nason (2011). The model confidence set. *Econometrica 79*, 453–497.

Harville, D. A. (2003). The selection of seeding of college basketball or football teams for postseason competition. *Journal of the American Statistical Association 98*, 17–27.

Harville, D. A. and M. H. Smith (1994). The home-court advantage: How large is it and does it vary from team to team. *The American Statistician 48*, 22–29.

Jones, M. B. (2007). Home advantage in the NBA as a game-long process. *Journal of Quantitative Analysis in Sports 3*(4), Article: 2.

Jones, M. B. (2008). A note on team-specific home advantage in the NBA. *Journal of Quantitative Analysis in Sports 4*(3), Article: 5.

Koopman, S. J. and R. Lit (2014). A dynamic bivariate poisson model for analyzing and forecasting match results in the english premier league. *Journal of the Royal Statistical Society A*, forthcoming.

Kubatko, J., D. Oliver, K. Pelton, and D. T. Rosenbaum (2007). A starting point for analyzing basketball statistics. *Journal of Quantitative Analysis in Sports 3*(3), Article: 1.

Loeffelhold, B., E. Bednar, and K. W. Bauer (2009). Predicting NBA games using neural networks. *Journal of Quantitative Analysis in Sports 5*(1), Article: 7.

Page, G. L., G. W. Fellingham, and C. S. Reese (2007). Using box-scores to determine a positions's contribution to winning basketball games. *Journal of Quantitative Analysis in Sports 3*(4), Article: 1.

Rosenfeld, J. W., J. I. Fisher, D. Adler, and C. Morris (2010). Predicting overtime with the pythagorean formula. *Journal of Quantitative Analysis in Sports 6*(2), Article: 1.

Schwertman, N. C., T. A. McCready, and L. Howard (1991). Probability models for the NCAA regional basketball tournaments. *The American Statistician 45*, 35–38.

Steckler, H. O., D. Sendor, and R. Verlander (2010). Issues in sports forecasting. *International Journal of Forecasting 26*, 606–621.

Stefani, R. T. (1977a). Football and basketball prediction using least squares. *IEEE Transactions on Systems, Man, and Cybernetics SMC-7*, 117–121.

Stefani, R. T. (1977b). Improved least squares football, basketball, and soccer predictions. *IEEE Transactions on Systems, Man, and Cybernetics SMC-7*, 117–121.

Stekler, H. O. and A. Klein (2012). Predicting the outcomes of NCAA basketball championship games. *Journal of Quantitative Analysis in Sports 8*(1), Article: 1.

Štrumbelj, E. (2014). On determining probability forecasts from betting odds. *International Journal of Forecasting 30*, 934–943.

Štrumbelj, E. and P. Vračar (2012). Simulating a basketball match with a homogeneous Markov model and forecasting the outcome. *International Journal of Forecasting 28*, 532–542.

Teramoto, M. and C. L. Cross (2010). Relative importance of performance factors in winning NBA games in regular season versus playoffs. *Journal of Quantitative Analysis in Sports 6*(3), Article: 2.

Timmermann, A. (2006). Forecast combinations. In *Handbook of Economic Forecasting*. Elsevier Press.