1 point = 1 point?

George A. Waters*
Department of Economics
Campus Box 4200
Illinois State University
Normal, IL 61790-4200

May 26, 2022

Abstract

Basketball coaches usually take out a star player in foul trouble due to the belief that games are often decided in the final minutes. Such a belief depends on the existence of the "rubber band effect," which means teams with the lead tend to relax so the lead reverts to zero, and implies that the lead, or margin, data is stationary. A theoretical model demonstrates the rubber band effect. A proper test of the null of a unit root in the margin data of a game allows for a random autoregressive coefficient as the alternative hypothesis. Such tests show evidence of the rubber band effect in at most a quarter of the NBA games from the 2020-20201 regular season. Tests on the aggregate data show no evidence of the rubber band effect.

Keywords: NBA, margin data, stochastic autoregressive model, RCAR, rubber band effect
JEL codes: C22

*gawater@ilstu.edu

In basketball is a point scored in the first minute worth as much as a point scored in the last second? You would think so since the winner is the team with the most points regardless of when they are scored. But I don't think so...

- Jeffrey Ely

I vehemently disagree with all of this thread, but otherwise sensible people such as @bcmassey take this view. (He and I have agreed to disagree). My opinion: one point = one point. This is why the usual foul trouble strategy is stupid. @dmorey can referee.

- Richard Thaler

Rubber band effect is one unique reason to the NBA that I can say with 100% certainty that a point is not a point. There may be others but to @R_ Thaler 's point, usually people think there are other reasons and they are usually wrong.

- Daryl Morey[1]

1 Introduction

When a star basketball player gets into foul trouble early in a game, in the vast majority of cases, coaches substitute for the star to maximize the player's availability at the end of the game. Such a decision is rooted in the belief that most games are decided in the last few minutes. Richard Thaler (quoted above) disagrees saying "one point equals one point," meaning a point scored in the last few minutes counts as much as a point scored earlier in the game.

Daryl Morey (quoted above) describes the *rubber band effect* as a reason points at the end of the game do matter more. If the team with the lead relaxes and its performance declines, the lead tends to shrink. Similarly, a team that is behind may expend more effort. In either case the lead tends to revert toward zero, and the score is closer at the end of the game with the rubber band effect. Hence, points at the end of the game do matter more, $1pt. \neq 1pt.$, and coaches have incentive to make sure their best players are available[2].

A simple two team, two shot model, where the lead could affect shooting probabilities demonstrates the rubber band effect. A made shot in the first period implies a lower expected score in the second, so it does matter less than a made shot in the second period. In an extension of the model to multiple periods, the margin is mean reverting and stationary under the rubber band effect, but it has a unit root if shooting probabilities are independent of the margin.

The proper test of a unit root in the margin data allows for the alternative to be a random coefficient autoregressive process. Using the test of Disasto (2008), a unit root in the margin data is rejected at a 10% significance level for less than a quarter of the games from the 2020-2021 NBA regular season.

[1] Twitter exchange on April 17,2021. https://twitter.com/dmorey/status/1383535982981517314

[2] Interestingly, Dr. Thaler does not believe this behavioral interpretation.

The test on the aggregate data for all 1080 games shows no correlation between shooting probabilities and the margin. Similar tests on turnovers, which are thought to measure player effort, show slightly less turnovers with a lead. There is no evidence of the rubber band effect in the aggregate data.

The work of Polson and Stern (2017), which builds on Stern (1994) is the most closely related. They model the margin as a continuous time Brownian motion process with a deterministic time trend, which is equivalent to a random walk with drift. They use Black Scholes formulas from finance to model win probabilities and their volatility. This model assumes scoring is an independent event, so it does not have implications for the question at hand.

The inclusion of the deterministic trend reflects the idea that the better team will tend to increase its lead over the game. Using an estimated trend changes the results dramatically, and the rubber band effect is much more apparent. However, using pre-game point spreads from betting site is a more reliable method, and the hypothesis of a unit root in the margin data is rejected for only a few more games than without any trend.

Section 2 describes a two period model demonstrating the rubber band effect. Section 3 extends that model to one with multiple periods that can describe margin data, and demontrates the need for a random coefficient version of the model, then gives results of the associated empirical tests. Section 4 reports results on the aggregate data, and Section 5 concludes.

2 The Contest

The rubber-band effect is apparent in a simple model of a contest between teams 1 and 2 who take two alternating shots in time $t = 1, 2$ that give one point with probability $p_{k,t}$ for team k and zero points for a missed shot. The score for team k in period t is $s_{k,t,}$, which is the the sum of the points $S_{k,t} = 0, 1$ achieved in time t and before such that $s_{k,t} = \sum_{h=1}^{t} S_{k,h}$. The margin m_t in time t is the difference in the scores

$$m_t = s_{1,t} - s_{2,t}.$$

If team k is in the lead entering period t, its players tend to relax so its probability of making a shot that period is αp_k where the parameter $\alpha \in (0, 1]$.

Consider the outcome where team 1 makes the first shot and misses the second, while the team 2 missed the first but makes the second, so the sequence of four shots is

$$(S_{1,1}, S_{2,1}, S_{1,2}, S_{2,2}) = (1, 0, 0, 1) .$$

The probability P of the outcome $(1,0,0,1)$ is

$$P(1,0,0,1) = p_{1,1}(1-p_{2,1})(1-\alpha p_{1,2})\,p_{2,2}.$$

The parameter α appears since team 1 had a lead after the shots in the first possession. The margin at the end of the contest is $m(1,0,0,1)=0$. Hence, the unconditional expected margin $E(m_2)$ is the sum over all 16 possibilities.

$$E(m_2) = \sum_{g=0}^{1}\sum_{h=0}^{1}\sum_{i=0}^{1}\sum_{j=0}^{1} P(g,h,i,j)\,m(g,h,i,j)\,.$$

The following proposition describes a case where a coach can make a decision that improves the probability of scoring in either the first period or second period, like the decision to take a star player out of the first half of a game to make sure that player can play late in the game.

Proposition 1 *If both teams have the same shooting probability such that $p_{k,t}=p$, then the difference Δ in an incremental increase in the shooting probability between the second and first shots for team 1 is*

$$\Delta = \frac{d}{dp_{1,2}}E(m_2) - \frac{d}{dp_{1,1}}E(m_2) = (1-\alpha)\,p^2.$$

Proof. See Appendix ■

Hence, if all shots are independent events, so that $\alpha = 1$, then there's no difference between making an improvement in $t=1$ or 2, as $\Delta = 0$. However, if the lead does matter and $\alpha < 1$, then there is a reason to take out the star in the first period, $\Delta > 0$. Points in time 1 matter less since they imply a lower expected score in time 2, the *rubber band effect* in Morey terminology.

The fact that the difference Δ depends on the shooting probability p, gives one reason that the issue of the rubber band effect is more of a concern in basketball, compared to other sports. In soccer (football...whatever) or hockey shooting percentages are much lower, so, even if the rubber band effect is present $\alpha < 1$, the low probability p means that it is quantitatively unimportant.

3 Empirical Model

If shots are independent events, the margin data has a unit root. However, testing the hypothesis that $\alpha = 1$ with game data presents some complications not handled by standard tests.

The process for the margin in the model above can be written

$$m_2 = m_1 + S_{1,2} - S_{2,2},$$

so extending the model above to n periods yields

$$m_t = m_{t-1} + e_t. \tag{1}$$

To model a basketball game, the stochastic term e_t should reflect the probabilities of scoring a free throw p_{ft}, a two-point shot p_{2pt} and a three-point shot p_{3pt}. Events such as steals and offensive rebounds are relatively rare and difficult to model explicitly and are excluded here. The stochastic term is as follows.

$$\begin{aligned} e_t &= \alpha \widetilde{S}_t - (1-\alpha)\,\widetilde{S}_{2,t} \quad \text{if } m_{t-1} > 0 \\ &= \alpha \widetilde{S}_t + (1-\alpha)\,\widetilde{S}_{1,t} \quad \text{if } m_{t-1} < 0 \\ &= \widetilde{S}_t \quad \text{if } m_{t-1} = 0 \end{aligned} \tag{2}$$

where the score of a team for a single possession is given by the distribution

$$\widetilde{S}_{k,t} = \left\{ \begin{array}{c} 1 \text{ with prob } p_{ft} \\ 2 \text{ with prob } p_{2pt} \\ 3 \text{ with prob } p_{3pt} \end{array} \right\} \tag{3}$$

and the aggregate score for a single possession is

$$\widetilde{S}_t = \widetilde{S}_{1,t} - \widetilde{S}_{2,t}.$$

The following lemma states that, for $\alpha < 1$, the time between crossings of zero for m_t is finite.

Lemma 2 *Assume $\alpha < 1$ in the process for m_t given by (1), (2) and (3). For any value of $m_t > 0$, the expected number of periods h such that $m_{t+h} < 0$ is finite. For any value of $m_t < 0$, the expected number of periods h such that $m_{t+h} > 0$ is finite.*

Proof. If $m_t > 0$ then the expected value of the next innovations is

$$E_t e_{t+1} = \alpha E_t \widetilde{S}_{1,t+j} - E_t \widetilde{S}_{2,t+j}$$

and so

$$E_t e_{t+1} = (\alpha - 1)\, E_t \widetilde{S}_{k,t+j} < -c,$$

for some constant $c > 0$, since $\widetilde{S}_{k,t+j} > 0$ and has the same distribution for $k = 1, 2$. Hence, there exists a positive integer h such that $E_t m_{t+h} = m_t - ch < 0$. The argument when $m_t < 0$ is analogous. ■

Corollary 3 *For $\alpha < 1$, the process for m_t does not have a unit root, is mean reverting and has finite variance.*

Proof. See Choi (2015) who relies on Engle and Granger (1987), which is based on Granger and Newbold (1977). ■

Corollary 4 *For $\alpha < 1$, the process for m_t is covariance stationary.*

Proof. Since m_t has finite variance, and the autocorrelation function $acf(\tau) = E(m_t m_{t+\tau})$ does not depend on t, it is covariance stationary. ■

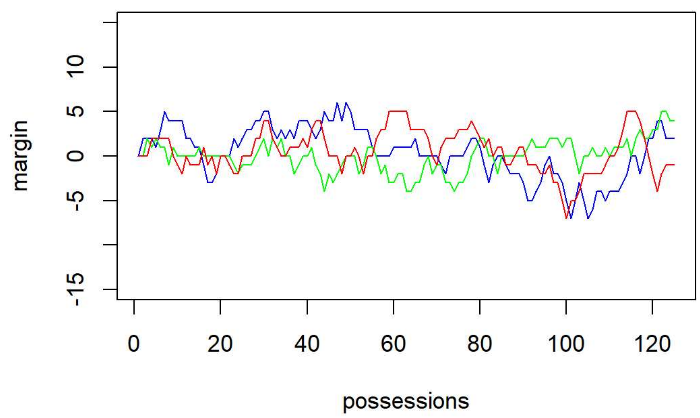

Figure 1

Simulation of m_t for $\alpha = 0.7$

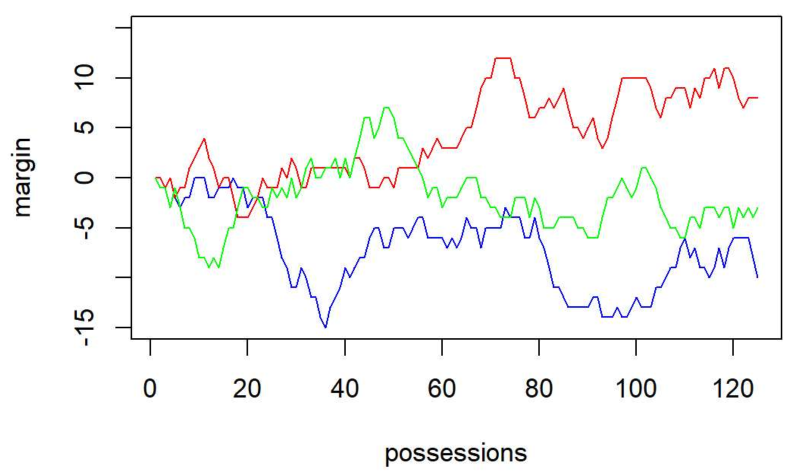

Figure 2

Simulation of m_t for $\alpha = 1.0$

Figures 1 and 2 show simulations of the model (1) and (2) for the value $\alpha = 0.7$, where the distribution of e_t depends on the margin, and the value $\alpha = 1.0$, where shooting percentages are independent of m_t. The probabilities are determined by the aggregate data from the regular season NBA games[3] in 2020-2021. The probabilities of taking and making a 3-point shot, a 2-point shot and a free throw are $p_{3pt} = 0.1211$, $p_{2pt} = 0.269$ and $p_{ft} = 0.145$. The series appear to be stationary in Figure 1, but the variances are increasing over time in Figure 2 reflecting a potential unit root.

3.1 Dickey-Fuller test for unit roots

The empirical model shows that tests for unit roots could provide a test of the independence of shooting percentages from the margin. However, there are some complications in the implementation.

The augmented Dickey-Fuller test is the standard method for testing for a unit root. Let the AR1 process for the variable y_t be $y_t = \rho y_{t-1} + \varepsilon_t$ where the parameter ρ is positive and the shock is distributed such that $\varepsilon_t \sim iid(0, \sigma_\varepsilon)$. For the equation $\Delta y_t = \theta y_{t-1} + \varepsilon_t$, the null of a Dickey-Fuller test is $H_0 : \widehat{\theta} = 0$, which is equivalent to a unit root process in y_t where $\rho = 1$. The test statistic of interest is the standard $t-$statistic from an OLS regression, but its distribution is non-standard.

Figure 4 shows the distribution of the $t-$statistic for an augmented Dickey-Fuller (ADF) test on m_t where four lags[4] of Δm_t are included in the regression to account for serial correlation.

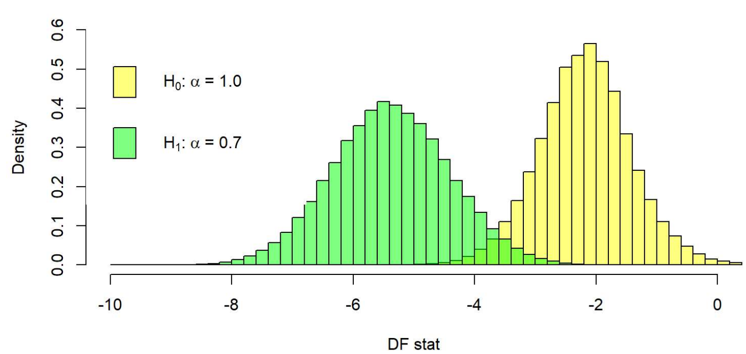

Figure 3

Estimated t-statistics for an ADF test on simulated data m_t using the model (1), (2) and (3).

[3] The source is the historical play-by-play data from bigdataball.com.

[4] Four lags was gave the best performance according to the SIC for most simulated series. Changing the number of lags does not make a qualitative difference. [check]

The distribution under the null of a unit root is typical of Dickey-Fuller results, and the distribution under the value $\alpha = 0.7$, shows that the test would reject a unit root for most of the simulations, so the ADF test works with a discrete distribution. The figure also confirms the intuition that the rubber band effect implies stationarity in the margin data.

The result shown in Figure 3 would bode well for inference on game data except that the sample size of 1250 possessions in the simulated data is 10 times greater than the sample size for a typical NBA game. For simulated data with 125 possessions, the distributions both look like the one under the null of a unit root and are almost indistinguishable.

Dickey-Fuller tests are not able to give evidence on the presence of unit roots in the margin data for NBA data. As is typical of test with the null of a unit root, Dickey-Fuller tests are biased against rejection. However, the econometric difficulties do not end here.

3.2 Stochastic coefficient model

One can write the stochastic term in the econometric model (2) more compactly using the term $\frac{m_{t-1}}{|m_{t-1}|}$ as an indicator function.

$$\begin{aligned} e_t &= \alpha \widetilde{S}_t + \frac{m_{t-1}}{|m_{t-1}|}(1-\alpha)\widetilde{S}_{k,t} \qquad \text{if } m_{t-1} \neq 0 \\ &= \widetilde{S}_t \qquad \text{if } m_{t-1} = 0 \end{aligned}$$

This formulation uses the fact the scoring distribution $\widetilde{S}_{k,t}$ is the same for both teams. The combination of the above representation of e_t with the auto-regressive process for m_t (1) can be written as follows.

$$\begin{aligned} m_t &= m_{t-1} + \widetilde{S}_t \qquad \text{if } m_{t-1} = 0 \\ &= \widetilde{\rho}_t m_{t-1} + \widetilde{S}_t \qquad \text{if } m_{t-1} \neq 0 \\ \text{for} \quad \widetilde{\rho}_t &= 1 - (1-\alpha)\frac{\widetilde{S}_{k,t}}{|m_{t-1}|} \end{aligned} \tag{4}$$

Remark 5 *The autoregressive coefficient $\widetilde{\rho}$ in (4) is stochastic for $\alpha < 1$.*

The representation confirms that the margin m_t is a unit root process for $\alpha = 1$. However, in the case of $\alpha < 1$, where shots depend on the margin, the margin process m_t is stationary according to Corollary 3, but the coefficient $\widetilde{\rho}_t$ on m_t is time varying and stochastic, since it depends on the variables $\widetilde{S}_{k,t}$ and m_{t-1}. Hence, we assume the coefficient is distributed such that $\rho_t \sim iid\left(\rho, \omega^2\right)$.

3.3 Time series results

Disasto (2008) develops an LM test for the null of a deterministic unit root $H_0 : \rho = 1, \omega = 0$ that allows for an alternative of a stationary process with a stochastic autoregressive coefficient $H_1 : \rho < 1, \omega \geq 0$.

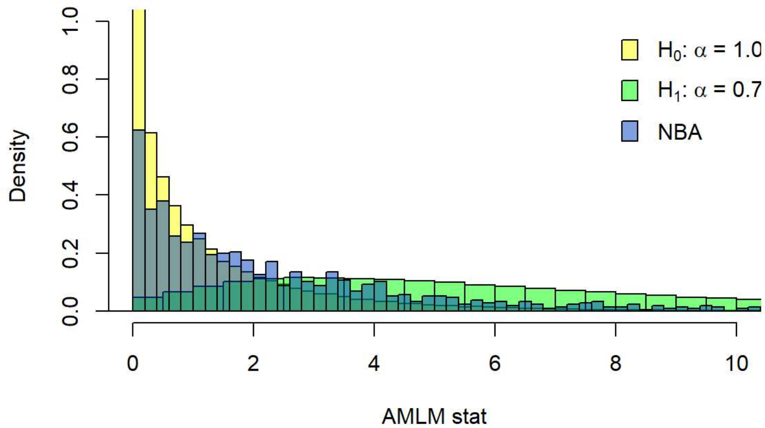

Figure 4

Distribution of estimates of the AMLM statistic of Disasto (2008) for simulated data m_t using the model (1), (2) and (3) under $\alpha = 0.7$ and $\alpha = 1.0$ with the estimated values of the AMLM statistic for the margin data from the games of the 2020-2021 games of the NBA season.

Figure 4 shows the densities of the Augmented Modified LM statistic[5] (AMLM) of Disasto (2008) for simulated data with a sample of 125 under the null[6] where $\alpha = 1$ and the alternative where $\alpha = 0.7$. The distributions are distinct, in contrast to the Dickey-Fuller tests, though there is considerable overlap so the size and power properties are not ideal. However, the simulations confirm that the AMLM statistic performs much better than standard tests in the presence of a random coefficient alternative, even with an error term with a discrete distribution.

The figure also shows the estimated AMLM statistics for margin data from the 1080 NBA games in 2020-2021. The distribution of the estimated data looks similar to that under the null of a unit root. To reject a unit root at a significance level of 5% (10%) requires the AMLM statistic to be greater than 4.71 (3.37), which occurs for 15.3% (24.5%) of the games[7]. So there are a few more rejections of the null of a

[5] See page 9 and equation (14) of Disasto (2008) for the precise formulation of the AMLM statistic.
[6] The graph is cropped for readability. The first visible bin bordering zero has an approximate height of 1.28.
[7] The results for the 91 playoff games was similiar at 15.4% and 20.9%, respectively.

unit root than would be expected due to chance, and the rubber band effect is apparent in a small fraction of games.

Since such unit root tests could be biased against rejection, an another important perspective on the dependence or shooting on the margin is a test with the null of stationarity with a stochastic auto-regressive parameter against the alternative of a deterministic unit root.

Trapani (2021b)[8] proposes just such a test in the context of the Random Coefficient Auto-regressive model (RCAR), of which model (4) is a special case. The proposed test statistic $\widehat{V}$ is a weighted average of the auto-covariances, see Appendix, of the variable Y_t, a transformation of the margin data

$$Y_t = \frac{a}{a + m_t},$$

for some value[9] of a. The test uses the intuition that the variance is constant under null hypothesis of a stationary process, so the statistic $\widehat{V}$ converges to a constant, but under the alternative of a unit root the variance of the margin increases with t so $\widehat{V} \to 0$.

The increasing variance under the unit root demonstrates the relationship of independence in shooting and a unit root in the margin data. Points near the end of a game matter more, since the score is usually close, when the team with the lead shoots worse. However, if shooting is independent of the margin, the lead at the end of the game tends to be the largest of the game.

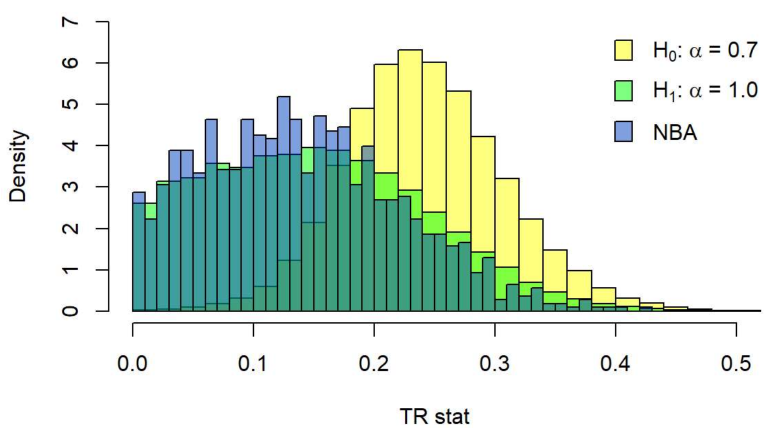

Figure 5

Distribution of the estimated $\widehat{V}$ for simulated data m_t using the model (1), (2) and (3) under $\alpha = 0.7$ and

[8] Trapani (2021a) presents a stationarity test for a general RCAR model, while Trapani (2021b) focuses on the AR1 case.
[9] Following Trapani (2021), the value is set to $a = 0.5$.

$\alpha = 1.0$ with the estimated values of $\widehat{V}$ for the margin data from the games of the 2020-2021 games of the NBA season.

Figure 5 shows the distribution of the estimated values of $\widehat{V}$ for 10,000 simulations of the margin data m_t using the model (1), (2) and (3). The distributions are distinct under the null of stationarity $H_0 : \alpha = 0.7$ and the alternative $H_1 : \alpha = 1.0$, with the latter clustering around zero as expected. However, the distributions are not sufficiently distinct to have much hope of rejecting independence with a reasonable combination of significance and power for an individual game.

Figure 5 also shows the same densities with the estimated Trapani (2021) statistic $\widehat{V}$ for the 1080 games from the 2020-2021 NBA regular season. The estimated density is quite similar to the one using simulated data under the hypothesis that shots are independent, $\alpha = 1$.

Using standard hypothesis testing, the null of stationarity is not rejected for the margin data for many games, but the power of the test is weak, a common problem with tests of stationarity, see section 4.7.9 of Choi (2015). For example, at a significance level of 5% (10%), stationarity is not rejected for 46.3% (37.7%) of the games, but the power of the test is .47 (.55).

The relationship of the distributions in Figure 5 is more compelling. At any reasonable level of significance, the Khomolgorov-Smirnov test rejects the hypothesis that the cumulative distribution function (cdf) of the statistics $\widehat{V}$ for the simulated data under the alternative is above that of the analogous cdf for the estimated from the NBA games. Hence, the density for the games is skews more toward zero than the density from the simulated data when $\alpha = 1$, which is evidence of independence of shooting from the margin.

3.4 Testing with a deterministic trend

To reflect the idea that the better team should increase its lead in expectation, Polson and Stern (2015) include a deterministic trend in their model of the margin for (American) football games. Following this intuition, the model in the present work would be written as follows.

$$\begin{aligned} m_t &= \mu t + m_{t-1} + \widetilde{S}_t \qquad \text{if } m_{t-1} = 0 \\ &= \mu t + \widetilde{\rho}_t m_{t-1} + \widetilde{S}_t \qquad \text{if } m_{t-1} \neq 0 \\ \text{for} \quad \widetilde{\rho}_t &= 1 - (1-\alpha)\frac{\widetilde{S}_{k,t}}{|m_{t-1}|} \end{aligned} \tag{5}$$

Presumably, players' performances are not affected by their knowledge of the expected trend, so the process for the margin is otherwise unchanged.

The next challenge is to determine an estimated of the trend coefficient $\widehat{\mu}$. One could simply estimate using OLS. Disasto (2008) recommends a preliminary estimate of $\widehat{\mu}$ assuming the null such that $\rho = 1$, then estimating the coefficient $\widehat{\rho}$. The process can be repeated to achieve convergence of the estimates, though repetition never changed the estimate significantly for the margin data. The AMLM statistic can then be constructed using the detrended data $m_t^* = m_t - \widehat{\mu}t$. Unlike with the Dickey-Fuller test, the distribution of the AMLM statistic testing against the null $H_{0:}\rho = 1$ is unaffected by the inclusion of the trend, according to Disasto (2008).

However, it is questionable whether an estimated parameter is the appropriate measure for the trend coefficient. The margin data includes all the idiosyncratic events of the game many of which are unrelated to the deterministic trend, which is summarized by the adage, "That's why they play the game."

Therefore, using the pre-game point spread from gambling sites, which is a forecast of the final margin for a game, is more appropriate, as done in Polson and Stern (2015). The estimate of the trend coefficient is the spread divided by the number of possessions.

Estimates of the AMLM statistic for the 1080 NBA games are computed using both the estimated trend and the spreads. Using the spreads, the results are relatively unaffected. At a significance of the 5% (10%), a unit root is rejected and the rubber band effect is apparent occurs in 17.5% (25%) of the games, which slightly more than the results without the trend[10].

Using the estimated trend, the results change dramatically. At a significance of the 5% (10%), a unit root is rejected for 48.9% (62.3%) of the games. As noted, these reliability of the results using the trend estimated ex-post is questionable.

Estimates of the test for stationarity from Trapani (2021b) using the trends show a similar pattern. Using the spreads, the distribution of the statistic $\widehat{V}$ for the NBA games is very similar to the one simulated under the alternative hypothesis $H_1 : \rho = 1$. Using the estimated trend, distribution shifts toward the one under the null of $H_0 : \rho = 0.7$. However, the power of the test is still low, and using the estimated trend is problematic.

3.5 Why not panel estimation?

Examining the distribution of estimates from a large number of time series looks like panel data. Panel estimation could improve the power of the tests if there is extra structure on the model. For example, one might focus on a specific null hypothesis such as all (of none) of the series contains a unit root, but those hypotheses are not particularly interesting here. Alternatively, games between the same teams, or that share another common factor, could be assumed to have the same estimate of $\widehat{\rho}$. However, such assumptions would

[10] At a 1% significance level, the critical AMLM value is 8.46, which is exceeded in 53 games or 4.9%.

be questionable, since there are many other aspects where the games would still differ. For example, even if the teams are the same, the players in those game could differ due to injuries.

Furthermore, the panel versions of the estimators of interest here, from Disasto (2008) and Trapani (2021a, 2021b), have not been developed. Hence, panel approaches are left for future research.

4 Tests on Aggregate Data

A more straightforward way to test for independence is to examine the relationship between shooting probabilities and margins across all the games. Naturally, including all 1080 games from the 2020-2021 season, which introduces many other sources of error, since the games are played at different times of the season between different teams on the different days of the week to name a few, but the larger sample should help.

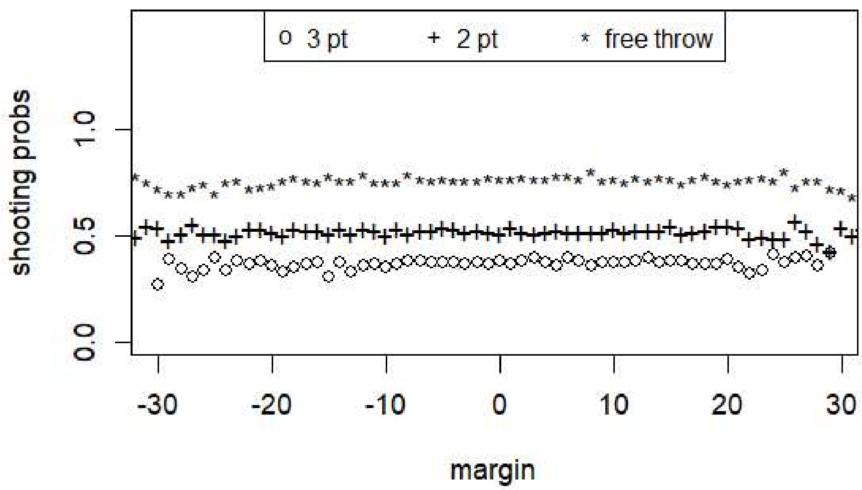

Figure 6

Figure 6 shows the probabilities[11] of making a 3-point shot, a 2-point shot and a free throw across varying margins or lead of the team shooting. If the rubber-band effect is present, probabilities should be lower with positive margins. However, there is no discernible pattern except for increased noise at the extremes, since large leads are less common. Such figures for playoff games were quite similar with more noise due to the smaller sample.

It is possible that player effort may be responsible for the rubber band effect, but that shooting probabilities are not the primary factor affected. Turnovers[12] are more likely to be committed by a team that relaxes with a lead and less likely for a team trying to catch up. Turnovers for different margins are distributed around a single peak at zero, since margins near zero are more likely.

[11] These probabilites of making a shot are commonly reported in basketball news are are different than the p's in the empirical model, which are the probability of taking and making a certain type of shot.

[12] Here, steals, balls lost out-of-bounds and shot clock violations are all counted as turnovers by the team with possesion.

To test this idea we construct a measure of turnovers per possession for varying margins, show in Figure 7 for margins between -40 and 40. The proxy for the number of possessions is constructed by restricting attention to turnovers and shots, while eliminating multiple free throw attempts[13]. Shots following offensive rebounds[14] are counted as an extra possession, which is reasonable, though not the common interpretation. Offensive rebounds are rare, so the decision has no quantitative implications. Both turnovers and possessions as functions of the margin are distributed around a peak at the margin of zero and slightly positively skewed. However, turnovers per possession is inversely correlated with the margin, see Figure 9. For the data shown, the correlation coefficient is -0.34, which is significantly less than zero with a p-value of 0.006. Turnovers provide no evidence of the rubber band effect.

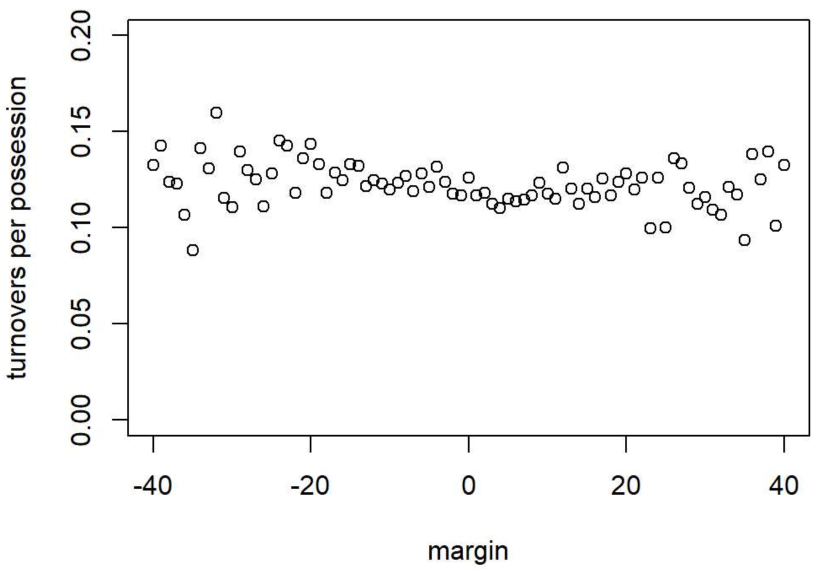

Figure 7

There is scope for further exploration of the data. The present work focuses on one season, but the NBA data used here is available back to the 2002-2003 season. College players are less experienced and may be more susceptible to psychological factors so the rubber band effect may be more apparent in data for those games. However, there are more mismatches where one team wins easily in college games, which might add to the noise in that data.

[13] Players can be awarded one, two or three free throws depending on the nature of the foul.
[14] An offensive rebound occurs when a team misses a shot but regains the ball for another shot opportunity.

5 Conclusion

Though Proposition 1 shows that Dr. Thaler's assertion that one point equals one point may lack nuance, the aggregate data confirms that shooting probabilities are independent of the lead and that taking out a star in foul trouble early in a game should not be the default strategy. The time series data offers modest evidence that the rubber band effect, when shooting does depend on the margin, is present in a small fraction of games, a quarter at most for any reasonable level of significance. The most reliable estimate accounts for a deterministic trend using pre-game spreads and a stochastic auto-regressive coefficient, and finds that in 17.5% of the games, an unit root is rejected at a significance level of 5% in the margin data and the rubber band effect is apparent.

The results have some relation to the "hot hand" hypothesis, which states that shooting probabilities rise for a player shooting after he or she has made a shot previously. Initial results[15] rejected the existence of a hot hand, which ran counter to the intuition of virtually all players, coaches and fans. More careful analysis such as Arkes (2010) has shown that hot hands do exist.

The two effects may also share an endogeneity issue in that the belief in their validity affects our ability to detect them in the data. Players believe in the hot hand, so when a player makes a shot, more effort is expended to defend that player lowering their shooting probability. Similarly, coaches believe the end of the game is more important so they conserve resources, like a star's playing time, with a lead and expend them when they are losing. If so, the margin data would appear to be stationary making the rubber band effect easier to detect. There is an endogeneity issue in both cases, but it makes the hot hand more difficult to detect in the data, but the rubber band effect should be more apparent.

There may be a cognitive bias behind the belief in the rubber band effect, which is held by virtually all players, coaches and fans, since the most memorable games are close and the result is determined by a few plays at the end. Furthermore, coaches are risk averse, and a loss due to a player fouling out gets more attention than a game where the team held a large lead for the whole game. There is limited statistical evidence for the rubber band effect, but it is doubtful that it is sufficient to sway coaches' beliefs.

References

Arkes, Jeremy (2010) Revisiting the Hot Hand Theory with Free Throw Data in a Multivariate Framework. *Journal of Quantitative Analysis in Sports* 6(1), https://doi.org/10.2202/1559-0410.1198

Bar-Eli M, Avugos S, Raab M (2006), Twenty years of "hot hand" research: Review and critique. *Psychology of Sport and Exercise* 7(6), 525-553.

Choi, In (2015) *Almost All About Unit Roots.* Cambridge University Press

[15] Gilovich et. al. (1985) is the seminal work. Bar-Eli et. al. (2006) survey the literature.

Disasto, Walter (2008) Testing for unit root processes in random coefficient autoregressive models. *Journal of Econometrics* 142, 581-609

Engle, Robert and Clive Granger (1987) Co-integration and Error Correction: Representation, Estimation, and Testing. *Econometrica* 55, 251-276.

Gilovich, T., Vallone, R., & Tversky, A. (1985). The hot hand in basketball: On the misperception of random sequences. *Cognitive Psychology* 17, 295–314.

Granger and Newbold (1977) *Forecasting Economic Time Series.* Academic Press, New York.

Polson, Nicholas and Hal Stern (2015) The implied volatility of a sports game. *J. Quant. Anal. Sports* 11(3): 145–153.

Stern, Hal (1994) A Brownian Motion Model for the Progress of Sports Scores. *Journal of the American Statistical Association* 89(427), 1128-1134

Trapani, Lorenzo (2021a) Testing for strict stationarity in a Random Coefficient AutoRegression. *Econometrics Reviews* 40, 220-256

Trapani, Lorenzo (2001b) A test for strict stationarity in a random coefficient autoregressive model of order 1. *Statistics and Probability Letters* 177, Article number 109164.

Appendix

The Proof of Proposition 1 follows.

Proof. Given shooting probabilities $p_{k,t}$ for team k in period t, let the probabilities for team 2 be the same such that $p_{21} = p_{22} = p_2$. The expected margin in the case of two periods and two teams is as follows.

$$\begin{aligned} E(m) &= p_{11}p_{12}p_2(1-p_2) - p_{11}(1-p_{12})p_2^2 + \alpha p_{11}p_{12}p_2(1-p_2) + 2\alpha p_{11}p_{12}(1-p_2)^2 \\ &\quad + p_{11}(1-ap_{12})(1-p_2)^2 - (1-p_{11})p_{12}ap_2^2 - 2(1-p_{11})(1-p_{12})ap_2^2 \\ &\quad - (1-p_{11})(1-p_{12})p_2(1-ap_2) + (1-p_{11})p_{12}(1-p_2)^2 - (1-p_{11})(1-p_{12})(1-p_2)p_2 \end{aligned}$$

Differentiation yields

$$\begin{aligned} \frac{d}{dp_{11}}E(m) &= p_{12}p_2(1-p_2) - (1-p_{12})p_2^2 + \alpha p_{12}p_2(1-p_2) + 2\alpha p_{12}(1-p_2)^2 \\ &\quad + (1-\alpha p_{12})(1-p_2)^2 + p_{12}\alpha p_2^2 + 2(1-p_{12})\alpha p_2^2 + (1-p_{12})p_2(1-\alpha p_2) \\ &\quad - p_{12}(1-p_2)^2 + (1-p_{12})(1-p_2)p_2 \end{aligned}$$

and

$$\begin{aligned}\frac{d}{dp_{12}}E(m) &= p_{11}p_2(1-p_2)+p_{11}p_2^2+\alpha p_{11}p_2(1-p_2)+2\alpha p_{11}(1-p_2)^2 \\ &\quad -\alpha p_{11}(1-p_2)^2-(1-p_{11})\alpha p_2^2+2(1-p_{11})\alpha p_2^2 \\ &\quad +(1-p_{11})p_2(1-\alpha p_2)+(1-p_{11})(1-p_2)^2+(1-p_{11})(1-p_2)p_2.\end{aligned}$$

Let the probabilities be the same so $p_{k,t}=p$ for $k,t=1,2$, where the above expression become

$$\begin{aligned}\frac{d}{dp_{11}}E(m) &= 3\alpha p^2(1-p)+2\alpha p(1-p)+(1-p)(1-\alpha p)+\alpha p^3 \\ \frac{d}{dp_{12}}E(m) &= p(1-p)+p^3+(1-p)^3\alpha p(1-p)^2+p(1-p)(1-\alpha p)\end{aligned}$$

The difference in the incremental changes in expected probabilities becomes

$$\begin{aligned}\frac{d}{dp_{12}}E(m)-\frac{d}{dp_{11}}E(m) &= p(1-p)(2+\alpha p)+p^3+(1-p)^3+\alpha p(1-p)^2 \\ &\quad -3\alpha p^2(1-p)-2\alpha p(1-p)^2-(1-p)(1-\alpha p)-\alpha p^3,\end{aligned}$$

which simplifies to the result in Proposition 1.

$$\frac{d}{dp_{12}}E(m)-\frac{d}{dp_{11}}E(m)=(1-\alpha)p^2.$$

■

The formulation for the test statistic $\widehat{V}$ of Trapani (2021b) is as follows.

The variable Y_t is the transformation of the margin data m_t

$$Y_t=\frac{a}{a+m_t}.$$

For a sample T, the mean of the variable Y_t is $\bar{Y}$, and the estimated autocoviances $\widehat{r}_j$ are given by

$$\widehat{r}_j=T^{-1}\sum_{t=j+1}^{T}\left(Y_t-\bar{Y}\right)\left(Y_{t+j}=\bar{Y}\right)$$

then the statistic $\widehat{V}$ is a weighted sum of the autocorrelations

$$\widehat{V}=\widehat{r}_0+2\sum_{j=1}^{H}\left(1-\frac{j}{H+1}\right)\widehat{r}_j.$$

The number of autocorrelations summed H depends on the sample size T such that $\frac{H(T)^3}{T} \rightarrow 0$ as $T \rightarrow \infty$. For sample size approximately $T = 125$, $H = 4$.