Counterpoints: Advanced Defensive Metrics for NBA Basketball

Alexander Franks*, Andrew Miller*, Luke Bornn, and Kirk Goldsberry

Harvard University,

Cambridge, MA, 02138

Email: afranks@fas.harvard.edu, acm@seas.harvard.edu

*These authors contributed equally to this work.

Abstract

Due to the ease of recording points, assists, and related goal-scoring statistics, the vast majority of advanced basketball metrics developed to date have focused on offensive production. It is straightforward to see who scored the most points in the 1985/86 season (Alex English, with 2414) or took the most 3-point shots in 1991/92 (Vernon Maxwell, with 473). However, try to look up who had the most points against in 2013/14, or who prevented the most shots from being taken that year, and the history books are, remarkably, empty. Thus we stand in a muddled state where offensive ability is naturally quantified with numerous directly-measured numbers, yet we attempt to explain defensive ability through statistics only loosely related to overall defensive ability, such as blocks and steals. Alternatively, we quote regression-based metrics such as adjusted plus/minus which give no insight into how or why a player is effective defensively. This paper bridges this gap, introducing a new suite of defensive metrics that aim to progress the field of basketball analytics by enriching the measurement of defensive play. Our results demonstrate that the combination of player tracking, statistical modeling, and visualization enable a far richer characterization of defense than has previously been possible. Our method, when combined with more traditional offensive statistics, provides a well-rounded summary of a player's contribution to the final outcome of a game.

1 Introduction

The two main objectives in the game of basketball are to score points on offense and to prevent points on defense. Unfortunately, to date, the vast majority of the game's analytics evaluate offensive performance while defensive performance continues to remain almost entirely overlooked. As a result, our ability to effectively assess overall basketball performance also remains significantly limited. Our research attempts to reduce this important limitation, leveraging an existing spatial regression model to propose new measures of defensive effectiveness.

The vast majority of contemporary basketball statistics are based on conveniently countable event types, but any reasonable defensive evaluation is largely incompatible with that approach [4,6,7,8]. While steals, blocks, and rebounds do provide some useful proxies for defensive skills, they represent small discrete signals within the perpetual broadcast of defensive play. Therefore, characterizations which rely on these event types are vulnerable to many forms of uncertainty - in short, such characterizations are unreliable. Fortunately, emerging forms of player tracking information curated by the National Basketball Association afford analysts an unprecedented opportunity to evaluate defensive performances in exciting new ways.

This paper introduces a new ensemble of defensive metrics designed to progress the field of basketball analytics by enriching the measurement of defensive play. We begin by describing a model that estimates defensive matchups for every moment in a basketball game. In other words, we estimate who is guarding whom at every moment of every NBA game during the 2013-14 season. For example, consider this "matchup box score" from the December 25th 2013 game between the Houston Rockets and the San Antonio Spurs. Here you can quickly identify who was guarding whom the most during that game.
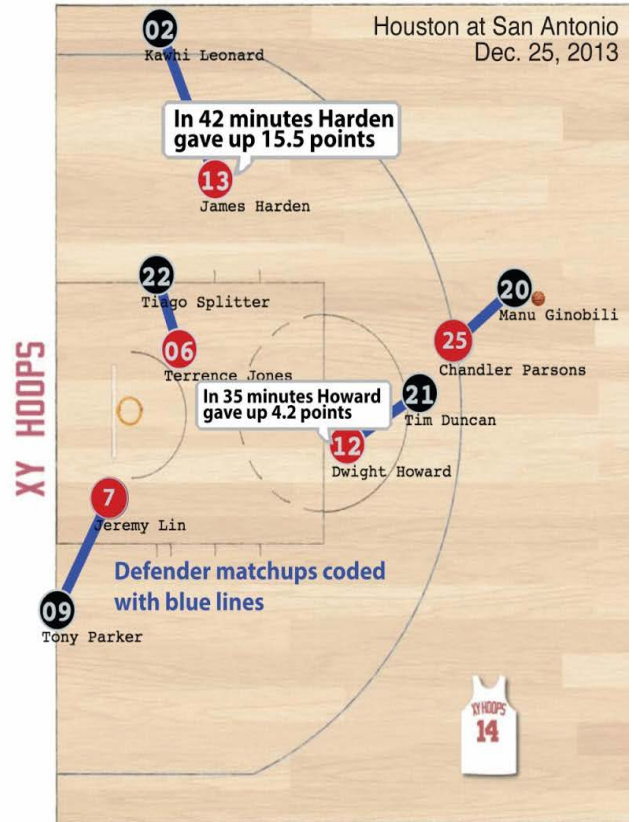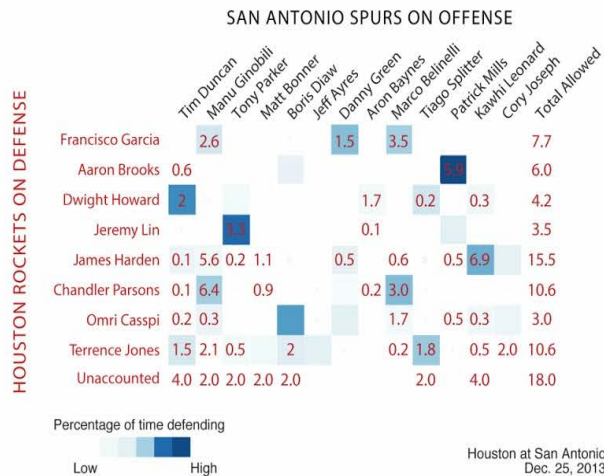
Figure 1. Matchup matrix for the Houston at San Antonio game on Dec 25, 2013. The matchup matrix has cells shaded according to the fraction of time spent guarding each offender. Counterpoints are assigned according to these fractions (see Methods). Points off of putbacks or fast breaks are not assigned to a defender ("unaccounted"). We visualize these responsibilities as a possession unfolds; the blue lines symbolize connections linking defenders to their offensive responsibilities (right side).

The matchup estimation process enables us to judge which defenders are responsible for which offenders at multiple stages of offensive possessions, including when points are scored. In other words, we know estimated defensive matchups for every offensive player who converted a field goal during the 2013-14 NBA season. Importantly, this knowledge is not limited to the moment at which the field goal was released; it pertains to the entire possession. This is key because the defender closest to the shot attempt is frequently not the most responsible defender.

In this paper, we leverage season-long matchup information alongside the spatial regression model of [2] to create a new ensemble of defensive metrics; we introduce five new artifacts:

1. **Volume Score:** The total magnitude of attempts which an individual defender faces.
2. **Disruption Score:** The degree to which an individual defender is able to reduce the effectiveness of his assignment's shots.
3. **Defensive Shot Charts:** Like shot charts, but for defensive play. Visual depictions of an individual's defensive prowess; we map both volume score and disruption score across the scoring area.
4. **Shots Against:** A weighted average of the shots attempted against the defender per 100 possessions.
5. **Counterpoints:** A weighted average of points scored against a particular defender per 100 possessions.

In the next section we introduce our methodological approach to measure these new metrics.

2 Methods

Who's guarding whom? Although the nuances of defensive play are difficult to analyze, there's no doubt that the performance of an individual player's defensive assignment is central to any assessment. Consequently, our analytical framework begins by estimating defensive matchups. In other words, we estimate who is guarding whom at any given moment. To identify this key information we estimate an average defender position as a function of offender location, ball location and the hoop location [2]. Mathematically, this means that the average location, μ, of player t at time k, is modelled as

$$\mu_{tk} = \gamma_O O_{tk} + \gamma_B B_t + \gamma_H H$$

with $\gamma_O + \gamma_B + \gamma_H = 1$, and O, B and H representing the offender, ball and hoop locations respectively. With these defensive centroids specified, we then use a hidden Markov model to express the evolution of defensive matchups over the course of the possession. Specifically, we model players' movement as a random walk around this evolving centroid. We infer the γ coefficients using weighted least squares combined with the expectation-maximization algorithm (see detailed methods in Appendix B).

Defensive Shot Charts, Volume Score and Disruption Score:

Using the defensive matchup model, we are able to take a much closer look at defensive skill and in particular how defenders affect shot selection and shot efficiency. In particular, we use the matchup model to define a defender "volume score" and defender "disruption score". The volume score quantifies how often a defender's matchup takes the shot when guarded by this particular defender. Conversely, the disruption score quantifies how much the defender reduces the opponents' shot efficiency. Although some have tried to address metrics to assess defensive volume ([3]), we are able define both the volume score and the disruption score in high resolution over regions of the court. In this paper, when computing all of the derived metrics, we exclude fast break possessions and possessions ending in fouls.
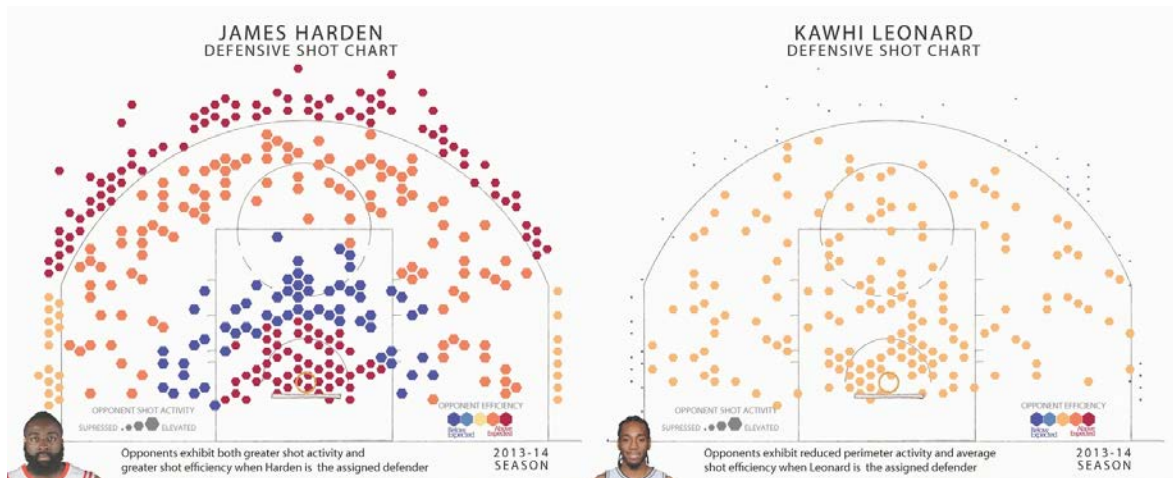

Figure 2a. Graphical depiction of a defender's volume (size) and disruption scores (color). Kawhi Leonard tends to suppress shots on the perimeter. More comparisons are provided in the Appendix.
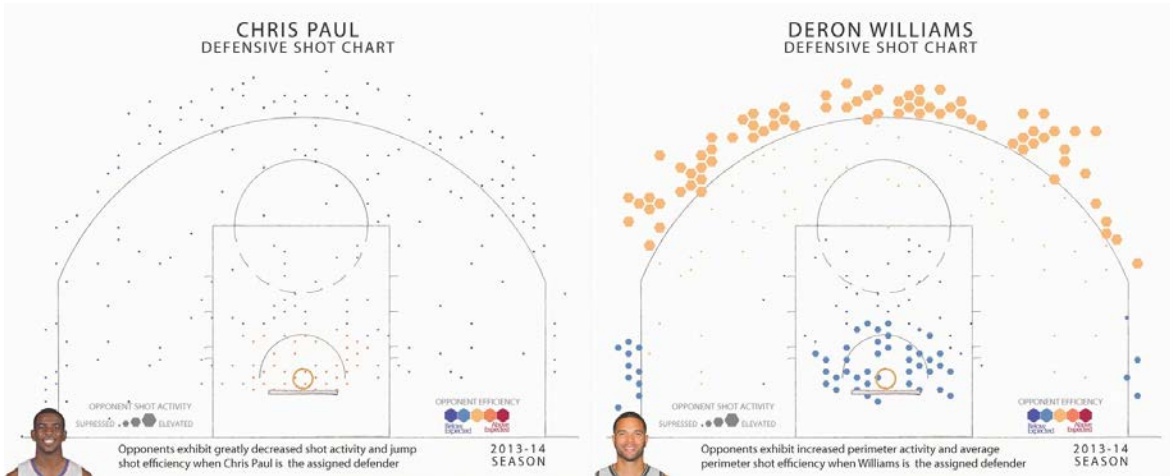
Figure 2b. Graphical depiction of a defender's volume (size) and disruption scores (color). Chris Paul has one of the best volume scores-- his matchup shoots against him significantly less often than expected, everywhere on the court. He also has the fewest average points against (Table 1).

To compute the volume score for the defensive shot charts, we follow the strategy of [2] and run a multinomial logistic regression, where each outcome represents a shooter-location pair. We include defensive matchups, and shooter and defender identities as predictors for the ultimate outcome (see Appendix B for details).

To compute disruption score for the shot charts, we use a logistic regression to predict makes and misses. Here, we use shooter and shot defender identities and defender distance to predict expected efficiency. The coefficients related to defender identities define the volume score and disruption score in each regression and correspond to the change in odds of a shot taken or shot made, respectively (see Appendix B for mathematical details). We graphically depict the volume score and disruption score through a new visualization we term the "defensive shot chart" (Figure 2).

Counterpoints. While volume and disruption scores (and hence the defensive shot chart) give insight into the what, why, and how of a player's defensive abilities, their primary limitation is that they are static-- they don't account for how the possession unfolds in time. Importantly, a defender usually does not guard the same offender for an entire possession. In these cases, a defender guarding the scorer at the beginning of the possession may be more or less responsible for the shot than a defender guarding the shooter one second before shot. We explore this notion by computing a variation on the defenders' volume score and disruption score at each moment t seconds before the time of the shot. Using this concept, we can identify how often a defenders initial matchup eventually shoots or scores.

First, we compute the "counter-attempts" for a variety of players. For every moment t seconds before the shot, we count the fraction of possessions at time t in which each defender is matched up against the eventual shot taker. We normalize the curve based on each defender's number of total possessions (see Appendix B) or the expected number of attempts each defender should face given the opponent's empirical shot frequencies. We compute the ratio of observed to expected shots against to rank defenders and define a time-varying volume score (Figure 2).

$$V_t = (Observed\ Attempts)_t/(Expected\ Attempts)_t$$

Where $(Expected\ Attempts)_t = \sum_i freq_{s_t(i)}$. Here, $freq_{s_t(i)}$ is the empirical shot frequency of the defenders' matchup at time t in possesion i. To understand this summary, we can imagine a completely average offensive team where each offensive player has a usage of .2. If defenders guarding this team never switch, they will always face .2 of all shots against, but their average original matchup score will be .2/.2 = 1. However, usages are distributed unequally (think Carmelo Anthony), and original matchup defensive attentiveness is also distributed unequally (think James Harden), so this ratio will deviate from one. A disruption score of less than 1 indicates defensive suppression of shots relative to average whereas a score over 1 indicates more volume than expected (Table 3).

Similarly, we compute another variant of the "disruption score" as the ratio of the observed number of points against (by matchup) to the expected number of points against (using player efficiencies)

$$D = (Observed\ Points\ Against)/(Expected\ Points\ Against)$$

Where $Expected\ Points\ Against = \sum_i (eff_{s(i)} * pts(i))$. Here, $S(i)$ is the shooter in possession i, $pts(i)$ is the point value of the shot in possession i, and $eff_{s(i)}$ is the empirical efficiency of the shooter. A disruption score of less than 1 is indicate of better than average shot disruption whereas a score of over 1 indicates poorer than average shot disruption.

Finally, we use this methodology to derive different metrics for points against, AKA "counterpoints". These metrics are appealing because their units are actual points scored. We define "counterpoints" in three different ways:

1) **Original matchup method::** Counterpoints are assigned to the defender who was guarding the shooter early in the possession (10-4 seconds before the shot occurs)
2) **Pre-shot matchup method:** Counterpoints are assigned to the defender who is guarding the shooter at the moment that the shot is taken.
3) **Fractional method**: Counterpoints are assigned proportionally. Each defender gets assigned points scored based on the fraction of the possession they spend guarding the scorer.

3 Results

The probability of a defenders' matchup taking the shot varies as we look back in time from the moment of the shot. For instance, intuitively, we expect centers to be the ones defending the shooter more often at the moment of the shot since they have a roll as rim protectors and help defenders. At the start of the possession however, the average defender would expect his matchup to be the one taking the shot roughly 20 percent of the time (1 in 5). The data clearly illustrate this phenomena. Roy Hibbert, arguably the best rim protector, defends nearly 45% of all shots taken (Table 1). However, between 8 to 10 seconds before the shot, his matchup is the shot taker between 20-25% of the time (see Appendix). In Figure 3, we illustrate the shot attempts curve for a selection of wing defenders.
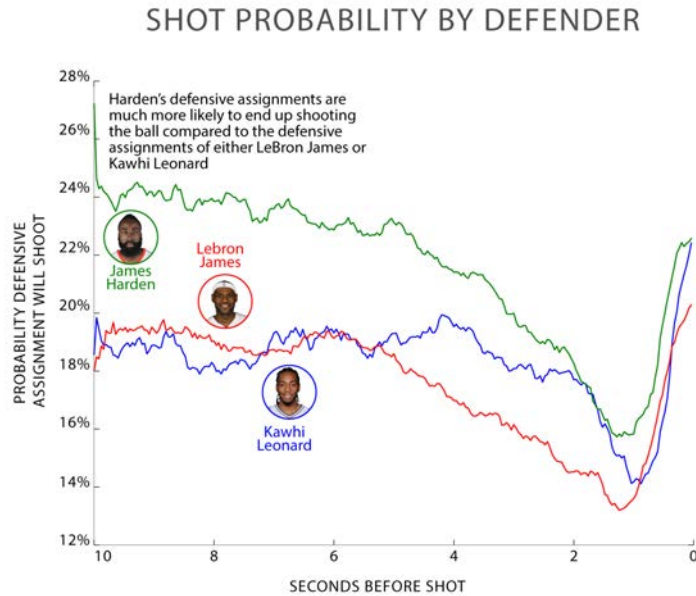

Fig 3: Shot attempts curves. Early possession matchup shoots against Harden more often than we would expect by chance. Immediately before the shot (4-1 second), defenders are less likely to be guarding the shooter. This is because the shooter is more likely to attempt a shot if he is momentarily unguarded in the second preceding the release of the ball.

Big Defenders: Shot Contests

	Most Contests			Least Contests	
Player	Percent Contest	Points Against	Player	Percent Contest	Points Against
1. Roy Hibbert	41. 9 %	16.6	1. David West	23.4 %	13.1
2. Robin Lopez	40.1 %	20.8	2. Mike Scott	23.9 %	15.0
3. Ian Mahinmi	39.3 %	16.4	3. Josh McRoberts	25.1 %	16.1
4. Joakim Noah	37.3 %	19.1	4. Blake Griffin	25.3 %	17.5
5. Timofey Mozgov	37.2 %	18.4	5. Jeremy Evans	25.6 %	15.8

Table 1: Top and bottom shot shot contesters. Roy Hibbert contests the shot on 41.9% of studied possessions. However, the fractional count of points against assigned to Hibbert is far less, which is indicative of the Pacer's defensive strategy.

Below, we present the the "original", "pre-shot", and "fractional" measurements derived from the "counterpoints" curve and the who's guarding whom model. Below we summarize three summaries of the "counterpoints" curve, and display values and rankings for a selection of back court players (Table 2).

Case Study 1: Original Matchup. The original matchup looks at the first can summarize defensive performance in the early possession regime by isolating seconds 10 through 4 before a shot attempt. For each moment before a shot, we examine two measurements - the average number of attempts taken and the average number of points scored against a defender.

Case Study 2: Before Shot Matchup
Another sensible summary of a defensive player's performance is to simply look at all match-ups for a defensive player at the moment of the shot (.5 to 0 seconds before release) and compute the "attempts against" and "points against" values. To do this, we look at all of the possessions played by a defender (where a shot occurs), and count how many times they were defending the shooter at the moment of the shot and how many points were scored per 100 possessions. However, these summaries must be interpreted with a grain of salt; rim protectors disproportionately face the shooter at the moment of the shot, thus inflating their "points against" score.

Case Study 3: Fractional Method
Both of these approaches yield interesting insights about defenders' strengths and weaknesses. However, both approaches have their flaws. The pre-shot matchup method disproportionately penalizes bigs who contests more shots. On the other hand, the early shot matchup method, may ignore the responsibility of the defender who contests the shot. We use the fraction method in Figure 1, to create the matchup matrix from the Rockets/Spurs game (12/25/2013).

Points Against Comparison (Back Court Defenders)

	Top Defenders				Bottom Defenders		
Player	Original	Shot	Fractional	Player	Original	Shot	Fractional
Chris Paul	14.4 (1)	17.7 (9)	10.8 (1)	Jrue Holiday	23.5 (61)	24.1 (50)	19.1 (63)
Norris Cole	15.0 (3)	17.0 (5)	11.1 (2)	Shaun Livingston	25.1 (63)	27.8 (62)	17.5 (62)
Nick Calathes	16.0 (5)	19.4 (18)	12.0 (3)	Jarrett Jack	21.1 (54)	22.3 (33)	17.5 (61)
C.J. Watson	18.8 (33)	19.3 (17)	12.0 (4)	Mo Williams	23.5 (62)	19.8 (19)	17.3 (60)
Greivis Vasquez	15.0 (2)	17.4 (7)	12.3 (5)	Patty Mills	23.1 (59)	23.1 (41)	17.1 (59)
Steph Curry	16.6 (7)	16.2 (2)	12.3 (6)	Kemba Walker	20.7 (51)	26.7 (60)	16.9 (58)

Table 2) Comparison of three points against metrics and their associated ranking for one defensive group (back court defenders). While highlighting slightly different aspects of defense, these metrics are largely consistent.

Points Against (Back Court Defenders)

	Top Defenders			Bottom Defenders	
Player	Volume Score (rank)	Points Against	Player	Volume Score (rank)	Points Against
1. Chris Paul	.79 (1)	10.8	1. Jrue Holiday	1.17 (61)	19.1
2. Norris Cole	.83 (3)	11.1	2. Shaun Livingston	1.11 (57)	17.5
3. Nick Calathes	.99 (39)	12.0	3. Jarrett Jack	1.05 (52)	17.5
4. C.J. Watson	.94 (19)	12.0	4. Mo Williams	1.12 (59)	17.3
5. Greivis Vasquez	.96 (27)	12.3	5. Patty Mills	1.18 (62)	17.1

Points Against (Wing Defenders)

	Top Defenders			Bottom Defenders	
Player	Volume Score (rank)	Points Against	Player	Volume Score (rank)	Points Against
1. Mike Dunleavy	.88 (18)	10.0	1. Jodie Meeks	1.07 (80)	18.5
2. Jordan Crawford	.90 (31)	10.9	2. Michael Kidd-Gilchrist	1.03 (71)	18.1
3. Eric Gordon	.82 (3)	12.2	3. Corey Brewer	1.04 (73)	18.0
4. DeMar DeRozan	.91 (35)	12.3	4. Evan Turner	.96 (47)	17.5
5. John Salmons	.88 (19)	12.4	5. James Harden	1.14 (87)	17.4

Table 3) Points against (fractional method) for back court defenders (top) and wing defenders (bottom). According to this metric, Chris Paul gives up the fewest points against per game on average and one of the lower volume scores (V_t). James Harden on the other hand, gives up the fifth most points against, largely because of the high volume of shots he faces relative to the expected number (1.14 times as much).

4 Discussion and Conclusion

Despite the player tracking revolution in basketball, assessments of performance have been heavily biased toward offensive play (e.g. [1, 5]). Examples of research addressing defense exist (see [3]) but are much less common. This research aims to correct this imbalance.. We presented a suite of metrics for measuring the defensive performance of NBA basketball players. Using optical tracking data and a model to infer defensive matchups at every moment throughout, we are able to provide novel summaries of defensive performance, and report "counterpoints" - points scored against a particular defender.

We believe that this is the first step toward a quantitative characterization of defensive culpability. Our methods allow us to look at the evolution of a possession and assign points against to particular defenders, explore how often opponents shoot against them, and *where* these defenders are most effective. When combined with offensive statistics, our method provides a well-rounded summary of a player's contribution to the final outcome of a game.

We explore three case studies of defensive summarization provide different lenses to view defensive performance. One key takeaway is that defensive *ability* is difficult to quantify with a single value. Summaries of points scored against and shots attempted against can say more about the team's defensive scheme (e.g. the Pacers funneling the ball toward Hibbert) than the individual player's defensive ability. However, we argue that these visual and statistical summaries provide a much richer set of measurements for a player's performance, particularly those that give us some notion of shot suppression early in the possession. For instance, it's not a statistical aberration that James Harden's early possession assignment tends to make a shot attempt at a much higher rate than Kawhi Leonard's assignment or that Chris Pauls has the lowest average points against per 100 possessions of any defender. We believe that the methodology presented sheds light on these contributions - contributions that have otherwise been left to vague and subjective judgements of skill. .

There are still many significant challenges in truly understanding defensive competence. Importantly, it is nearly impossible to assess defensive ability without understanding defensive intent. "Who's guarding whom" is only one way to understand intent. Without understanding team strategy, it is very difficult to know who a defender is *supposed* to be guarding or when they are supposed to help on defense. Future work on quantifying defensive ability ought to incorporate defensive schemes and intent into measurements of both the team and individual players. This can be

accomplished by combining knowledge from basketball experts with more complex (yet accurate) models of player interactions.

Nevertheless, our results demonstrate that the combination of player tracking and statistical modeling yield a far richer characterization of defense than has previously been possible.

5   References

[1] Cervone, Dan, et al. "POINTWISE: Predicting Points and Valuing Decisions in Real Time with NBA Optical Tracking Data." MIT Sloan Sports Analytics Conference 2014 (2014)

[2] Franks, Alexander, et al. "Characterizing the Spatial Structure of Defensive Skill in Professional Basketball." *To Appear in the Annals of Applied Statistics* (2014)

[3] Goldsberry, Kirk, and Eric Weiss. "The Dwight effect: A new ensemble of interior defense analytics for the NBA." MIT Sloan Sports Analytics Conference, 2013.

[4] Kubatko, Justin, et al. "A starting point for analyzing basketball statistics."*Journal of Quantitative Analysis in Sports* 3.3 (2007).

[5] Miller, Andrew, et al. "Factorized Point Process Intensities: A Spatial Analysis of Professional Basketball." *Proceedings of The 31st International Conference on Machine Learning*. 2014.

[6] Omidiran, Dapo. "A new look at adjusted plus/minus for basketball analysis." MIT Sloan Sports Analytics Conference 2011 (2011)

[7] Sampaio, Jaime, et al. "Discriminant analysis of game-related statistics between basketball guards, forwards and centres in three professional leagues."*European Journal of Sport Science* 6.3 (2006): 173-178.

[8] Sampaio, Jaime, Eric J. Drinkwater, and Nuno M. Leite. "Effects of season period, team quality, and playing time on basketball players' game-related statistics." *European Journal of Sport Science* 10.2 (2010): 141-149.