Journal of Quantitative Analysis in Sports

Volume 5, *Issue* 1	2009	*Article* 1

Importance of Teammate Fit: Frescoball Example

Dean Oliver, *Basketball on Paper*
Mike N. Fienen, *University of Wisconsin - Madison*

Recommended Citation:
Oliver, Dean and Fienen, Mike N. (2009) "Importance of Teammate Fit: Frescoball Example," *Journal of Quantitative Analysis in Sports*: Vol. 5: Iss. 1, Article 1.

DOI: 10.2202/1559-0410.1098

Importance of Teammate Fit: Frescoball Example

Dean Oliver and Mike N. Fienen

Abstract

The concept of teammates whose skills `fit' together is explored through frescoball, the beach paddleball game with two people hitting a ball back and forth trying to keep it in the air. This simple game can be analyzed through Markov Chains to obtain the expected number of hits between the two players as a function of two skills, which we label as `athleticism' and `consistency.' Using this theoretical model and conceptual values of parameters, we examine various combinations where the complementarity of skills between teammates enhances the performance of the team. Given the conceptual model of team performance commonly used in sports analysis – that the sum of player ability equals team performance – we look at how such a conceptual model mismatches team performance, leaving the difference as `fit' of teammates. We choose three examples to illustrate characteristics of fit, particularly where player marginal value varies depending on who they are paired with. Further, simulating 50-player leagues where player movement is simulated by keeping players together if the team is successful and moving them if not, we show how estimates of `fit' and player ability can be confounded. Ultimately, we seek to frame the discussion on teammate `fit,' for which there is no attempt to quantify in sports literature.

KEYWORDS: fresco, chemistry, fit, collaboration, player value, synergy

Author Notes: The authors wish to thank Mark Pankin for his initial insight, Scott Page and Mike Tamada for some useful discussion on background, and Marcia Fournier for introducing frescoball and tolerating all the scientific discussion of it.

1. Introduction

The concept of teammate "fit" is nearly unquestioned in team games. Though it is understood nearly without definition by most sports fans, we define "fit" as the complementary skill sets across teammates that create a team value different than the sum of the player values. To borrow from a common phrase, the whole may be greater than or less than the sum of the parts and that difference is what we call "fit."

"Fit" is implicitly considered in personnel decisions, by obtaining players at different positions in football, baseball, basketball, etc. – these positions bring different necessary skills to accomplishing team goals. Popular discussion of sports frequently refers to "fit" as an element of decision-making. For example, in the summer of 2007, *The Sporting News* wrote the following: "Because of his superior arm strength and size, [Daunte Culpepper] would be a better fit than [Joey] Harrington in Petrino's [Detroit Lions] offense." (Crossman, 2007)

At this point, we wish to distinguish "fit" from "chemistry", another word that is sometimes used interchangeably with fit. In particular, we are interested in *skill sets* that are complementary in a team sport. The term "chemistry" is sometimes used as a loose synonym for complementary skill sets, but it also is used to reflect how well teammates get along personally, in the implication that psychology has an effect on the underlying probabilities. There is some evidence in the literature that psychology does impact underlying performance as Totterdell (2000) showed a relationship between the mood of teammates and the performance of the team. However, we are not attempting to identify aspects of human psychology with this work. Rather, we are interested here solely in the complementary skill sets that we defined as "fit" above, completely bypassing psychology by working with synthetic players.

Although the word "fit" is so commonly used, there has been little work to quantify it explicitly. A few relevant efforts can be noted. First Bill James (2002) splits out the approximate credit between hitters (~48%), pitchers (~35%), and fielders (~17%) for the success of a baseball team. Similarly, Schatz (2006) credits offensive linemen and running backs in the cooperative act of running the football, giving varying credit depending on the overall result. In both of these cases, however, the "fit" of teammates is again implicit – hitters need pitchers and fielders who also need hitters, offensive linemen need running backs who need linemen. While a (poor) argument can be made that absent pitching, baseball teams would be 35% less successful, neither of these cases really quantifies the difference between the whole and the sum of the parts.

Beyond this, a recent paper on the NFL Draft (Fry, et al., 2007) uses dynamic programming and a simplification of "need-based" strategy to compare it to best-player-available draft strategy. That paper, however, uses a fantasy-league scoring mechanism and doesn't look explicitly at how the game structure of football leads to complementarity of teammates. Rather, it views "need" from a fantasy player's perspective where fantasy rules dictate having to draft a player at each of several positions. This perspective doesn't explicitly quantify fit in the real game of football.

Outside of sports, Page (2007) compiles a large series of peer-reviewed work on how diverse groups of people can be more successful at solving problems than less diverse but perhaps more talented groups. This suggests a "fit" effect in problem solving, but is not specifically related to sports.

It is our goal in this paper to assess how much the complementarity of skills enhances (or detracts) from the sum of player abilities in a simple game. In particular, we build a theoretical model of a real game called frescoball and use conceptual parameterizations to build a framework for discussing and quantifying "fit."

2. Frescoball Game Description

Frescoball (also known as beach paddleball, matkot, frescobol, or smashball) is a simple game of teamwork, involving two people, each with a paddle, and one ball (something like a racquetball or squash ball), standing about 15-30 ft apart hitting the ball back and forth to each other for as long as possible without the ball hitting the ground. Our discovery of the game was on the beaches of Brazil, where good players hit the ball hard and straight, keeping the ball alive for up to about a minute. Though score is not often kept, two contests can be envisioned. First, there is the extended game where the two players just try to keep the ball alive as long as possible. Second, there is the speed game where the two players maximize the number of hits within a given time frame (e.g., one minute). Though we focus here on the extended game, with either of these scoring rules, frescoball is conducive to a study on teammates collaborating with different skill sets.

The simplicity of frescoball – hitting a ball back and forth to a teammate – allows for a simple Markov model of the game, where the probability of successfully returning the ball is considered simply the result of the current state. This method follows analysis of other racquet games, for example, the work on table tennis by Noubary (2007). Unlike table tennis, frescoball is cooperative (not competitive like table tennis) and we add additional detail to the probabilities of hitting back and forth. Specifically, we define two skills, which we label as athleticism and consistency. An athletic person can dive or be agile to return shots that are harder to get a racquet on. A consistent player will return a well-hit shot to their teammate so that it is also hittable. We model this with two "zones" where shots get hit – Zone 1, which is an easy shot to hit back safely, and Zone 2, which is more difficult. Quantitatively, we define the "Athlete" as having a higher chance of returning a shot from Zone 2 than the "Non-athlete"

$$p_{A21} + p_{A22} > p_{N21} + p_{N22}, \tag{1}$$

where the A and N subscripts represent Athlete and Non-athlete, respectively, the first number represents the zone hit from, and the second number represents the zone hit to. The "Consistent" player has a higher chance of returning a shot in Zone 1 (an easy shot) than the "Inconsistent" player

$$p_{C11} + p_{C12} > p_{I11} + p_{I12}. \tag{2}$$

where the C and I subscripts represent "consistent" and "inconsistent," respectively. Other definitions or constraints are possible, but these are sufficient for the examples below. (There is no academic frescoball literature that we could find to support or add to these classifications, suggesting that we academic researchers have not spent enough time at the beach.)

There are five states of the game:

1. In Player 1's "sweet spot", where it is easy to hit back and control (Zone 1).
2. In Player 1's reach, but harder to hit (Zone 2) – called a "stretch" in this paper.
3. In Player 2's sweet spot.
4. In Player 2's stretch.
5. Ball out of play (game is over).

For simplicity, we assume that the serve to start the game is in one of the player's sweet spots – either State 1 or State 3 above. (This assumption could be relaxed in a variety of ways, but none would change the fundamental nature of the analysis below.) A game continues until State 5 is reached. Using the language of Markov Chains, State 5 is an absorbing state – once it is reached, it is not departed from and the game is over. The "state matrix" has a 1 in the corresponding zone where the ball is and zeroes elsewhere.

What determines how the game plays out are the probabilities of the two players hitting to the different zones, called the transition matrix in Markov Chain work. This matrix is shown below:

$$\mathbf{T} = \left(\begin{array}{cccc|c} 0 & 0 & p_{113} & p_{114} & p_{115} \\ 0 & 0 & p_{123} & p_{124} & p_{125} \\ p_{211} & p_{212} & 0 & 0 & p_{215} \\ p_{221} & p_{222} & 0 & 0 & p_{225} \\ 0 & 0 & 0 & 0 & 1 \end{array}\right) \tag{3}$$

where p_{ijk} represents the probability of player i hitting the ball from zone j to state k. The state matrix left multiplies the transition matrix to obtain the probabilities for the subsequent state. The first row reflects that if we are in State 1 or State 2, there is a 0% chance of moving into State 1 or State 2 since Player 1 cannot hit it to himself. The ball will be in either State 3 with Player 2 having it in his sweet spot (with probability p_{113}), State 4 with Player 2 having to stretch (with probability p_{114}), or State 5 with the point being over (with probability $p_{115} = 1 - p_{113} - p_{114}$).

Following the methodology of Noubary, we partition the transition matrix as

$$\mathbf{T} = \begin{pmatrix} \mathbf{Q} & \mathbf{R} \\ \mathbf{0} & 1 \end{pmatrix} \tag{4}$$

where $\mathbf{Q}$ is the 4×4 submatrix in the upper left of eq. 3 that represents the transition among states 1 through 4 (i.e., continuous play), $\mathbf{R}$ is the 4×1 vector in the upper right (designating the probabilities to end the game), $\mathbf{0}$ is the 1×4 vector of zeroes in the lower left, and 1 is simply the number 1 in the lower right. The long-term steady-state defined by this transition matrix is purely State 5; that is, the game will end eventually.

The success of a team of two players is represented here by how many total hits it will take (summing each of the two players) to reach State 5. This number can be found by summing the rows of the "Fundamental Matrix", $\mathbf{F} = (\mathbf{I} - \mathbf{Q})^{-1}$. Specifically, the sum of row 1 yields the expected number of hits from state 1 to the end of the game, the sum of row 2 yields expected number of hits from state 2 to the end of the game, etc. By assuming that the serve is represented by either State 1 or State 3, the sums of these two rows are the only ones of interest for this exercise. Inverting the 4×4 Fundamental Matrix is complicated and is thus shown in the Appendix A. The analytical results of eqs. A16-A19 in Appendix A will be used in the subsequent analysis.

3. Synthetic Teams

With the above method for estimating the expected remaining hits, we want to create some synthetic teams to illustrate characteristics of “fit.” Equations 1 and 2 structurally define four types of players:

- Athletic and consistent (designated herein as “AC”).
- Athletic and inconsistent (AI).
- Non-athletic and consistent (NC).
- Non-athletic and inconsistent (NI).

A team of two ACs will likely be the best. A team of two NIs will likely be the worst, but other teams are indeterminate in quality. The examples that follow will look at all 10 possible teams for each parameterization of the probabilities: AC with AC, AI, NC, and NI; AI with AI, NC, and NI; NC with NC and NI; and NI with NI. (We note here that the labels of athleticism and consistency are used here more for convenience in describing the game than for arriving at general rules for how to combine players.)

The players' characteristics used[1] in all three examples are shown in Table 1. In the first two examples, the “consistent” characteristics, p_{11} and p_{12}, are the same for AC and NC players. The “athletic” characteristics, p_{21} and p_{22}, are the same for AC and AI. The “inconsistent” characteristics, p_{11} and p_{12} again, are the same for AI and NI. Finally, the “non-athletic” characteristics, p_{21} and p_{22}, are the same for NC and NI. Throughout all examples, we maintain two restrictions on the probabilities:

$$p_{11} \geq p_{21} \tag{5}$$

which means that the chance of hitting to the sweet spot is at least as high when hitting from the sweet spot as from the stretch, and

$$p_{11} + p_{12} \geq p_{21} + p_{22} \tag{6}$$

which means that the chance of returning safely are, in a mean sense, at least as high hitting from the sweet spot as from the stretch.

[1] This analysis assumes constant probabilities of hitting the ball in the different zones for all players. Players legitimately do learn over time and from good coaches to improve their skills. They also get older and skills deteriorate. And there can be a period of adjustment to teammates where probabilities are different. These effects were not considered.

Table 1
Characteristics of Players in the Examples

	Example 1				Example 2				Example 3			
	AC	AI	NC	NI	AC	AI	NC	NI	AC	AI	NC	NI
p_{11}	70%	40%	70%	40%	80%	40%	80%	40%	60%	40%	70%	40%
p_{12}	25%	50%	25%	50%	17%	40%	17%	40%	35%	40%	25%	40%
p_{13}	5%	10%	5%	10%	3%	20%	3%	20%	5%	20%	5%	20%
p_{21}	30%	30%	10%	10%	40%	40%	10%	10%	30%	40%	50%	20%
p_{22}	40%	40%	40%	40%	40%	40%	50%	50%	50%	30%	10%	40%
p_{23}	30%	30%	50%	50%	20%	20%	40%	40%	20%	30%	40%	40%

With these probabilities, we find average hit count results for the teams using the equations in Appendix A. We don't yet label anything as the "team score" because there are two possible average scores for each team, depending on which of the players serves (which means starting in state 1 or state 3). For example, on the team with AC and NC in example 1 (abbreviated as AC/NC), if the serve is by AC, the team's score would be 6.5; but if the serve were by NC, the score would be 6.6. (Because these are results based on theory, there is no question that noise is affecting these numbers and comparisons of numbers that appear to be similar are reasonable.) Who does the serving matters in the results of the team, though it is a small difference in this example. For a team of AC/AI in the same example, AC serving gives a benefit of 0.5 over AI serving, so it can matter more. As we look at the relative performance of teams, we will assume that teams understand who is better to be serving and always start with that person; e.g., the team of AC/NC is considered to have an average score of 6.6, not 6.5. (Subsequent analysis could be done using the average of the server scores, but it would not provide additional insight.) Hence, future reference in this document to the "team score" or "team average score" refers to the maximum of the sum of the rows of eqs. A16-A19.

These examples are meant to illustrate various characteristics of "fit." Although "fit" is defined above as the difference between the team results and the sum of the parts, this definition implies various characteristics that can be easily seen through synthetic teams.

Example 1

In this first example, the athletes are quite a bit better than the non-athletes at returning a tough shot and the consistent players are a lot better than inconsistent players at returning to Zone 1.

The way we will present the results of the ten possible teams is in Table 2. This shows the average score for a player on the left paired with a player at the top. The third row first column shows 6.6 as the score for AC/NC, as does the first row third column, since it is the same team. This presentation is useful to rank each of the teams in a column, as the top is always AC with another teammate and probably best, with the bottom always NI with a teammate and probably worst of the column.

Column 1 of Table 2 shows that AC/AC is the best team with an average score of 8.1 hits. The second best team in that column is AC/NC – two consistent players. Then it is AC/AI and AI/NI last. Replacing AI with NC as a teammate for AC gives a positive benefit to the team. Column 2 of Table 2 shows AC/AI as the best team with an average score of 6.3. The second best team is AI/AI at 5.2, followed by AI/NC and AI/NI. In this case, replacing AI with NC as a teammate gives a negative benefit.

This is our first characteristic of "fit" – that replacing one player with a specific other player in one situation makes the team worse, but makes it better in another. One could say that NC "fits" better with AC and NC, but AI "fits" better with AI and NI. NC is consistent enough that he doesn't need the extra athleticism of AI because AI is too erratic – "fit" is a more concise way to say that. Columns 3 and 4 show similar indications of fit, where NC is better with NC as a teammate over AI, but where NI is better with AI as a teammate over NC.

This characteristic can be used to explain trades between teams that are said to help both teams. In this case, two AI/NC teams would both do well to trade between each other to arrive at one AI/AI team and one NC/NC team. The AI/NC teams had average hit totals of 5.1, whereas the new teams are at 5.2 and 5.5, respectively. Ordinal rankings may decline for the new AI/AI team, which may be all that matters in a zero-sum league defined by ordering, but both improve their average hit totals.

This "fit" effect can be found with various parameterizations of the probabilities, but definitely not all of them. It is possible to make AI or NC "good" enough relative to the other that fit is not enough to overcome that "ability". We place "good" and "ability" in quotes because we have not defined *individual* value yet, but both are at least intuitively understood by most readers. Another way to express this idea is that individual value can appear to be a function of context. We will discuss in a later section a measure of individual value and the resulting way to measure fit.

Table 2
Results of Example 1: Average Scores for Different Player Combinations

	AC	AI	NC	NI
AC	8.1	6.3	6.6	5.2
AI	6.3	5.2	5.1	4.4
NC	6.6	5.1	5.5	4.3
NI	5.2	4.4	4.3	3.5

Example 2

This example, whose parameters are shown in Table 1, has results shown in Table 3.

This example has all of the columns ranked the same way, all having the best team atop, the worst team on the bottom, with the second best team in the third row. Unlike the last example, where some reordering occurred, this example shows how the marginal value of a player is strongly affected by context, another characteristic of fit (and, really, just a generalization of the characteristic in Example 1). For example a team of NI/NI has an average score of 3.5, the worst team of all. Replacing one of these players with AC gives the team a score of 5.8, a marginal increase of 2.3. The apparent marginal value of AC relative to NI is then 2.3 hits. But if you look at the AC/NI team and replace the NI with another AC, the team jumps from 5.8 to 14.8! The exact same replacement has a marginal value of 9.0 hits, more than three times the marginal rate of replacing the first NI. This variable marginal value of a player as a function of teammate is a clear characteristic of "fit."

This is also (one reason) why barroom discussions of player value can get so heated – maybe not of frescoball players, but of soccer players, basketball players, etc. Teams can change hugely or only a small amount with the exact same player added. Context can make the marginal value of a player look a lot better or a lot worse. So anecdotal evidence – that of barroom arguments – can be very misleading. In the real world, however, we often don't have a lot of contexts to compare to in order to get a better sense of individual player value, a theme we return to later. In this theoretical model, we can look at all contexts.

Table 3
Results of Example 2: Average Scores for Different Player Combinations

	AC	AI	NC	NI
AC	14.8	7.0	10.9	5.8
AI	7.0	5.0	5.7	4.3
NC	10.9	5.7	8.1	4.7
NI	5.8	4.3	4.7	3.5

Example 3

A third example illustrates an interesting situation. The probabilities for the different players are shown again in Table 1. The results are in Table 4. In this case, the most interesting column is the third column – the teams with the NC player on them. In particular, note that a team of NC/NC actually is better than a team of NC/AC! This is despite the fact that teams with AC win each of the other columns and AC is designed to be the best overall.

Part of the reason this could happen is because the parameters varied more within the definition of "athletic." Players 1 and 2 have greater totals of $p_{21} + p_{22}$, but players 3 and 4 have greater p_{21}. Essentially, this illustrates another skill – athletic ability to hit to the sweet spot. This could occur with consistency as well, breaking it down into overall consistency and consistency to hit back to the sweet spot. By having more variation, there is essentially more specialization, allowing fit to become more important. Though some may consider this "cheating" to obtain such an odd result, we do this to illustrate the occasional translational difficulties between the words used to describe players and numerical quantifications. Sometimes examples like this allow us to redefine words like "athletic" to be more specific and avoid this kind of cherry-picking of results.

Table 4
Results of Example 3: Average Scores for Different Player Combinations

	AC	AI	NC	NI
AC	8.9	6.0	7.8	5.2
AI	6.0	4.2	5.3	4.0
NC	7.8	5.3	7.9	4.9
NI	5.2	4.0	4.9	3.6

Further, this suggests to us that more freedom to define skills may cause "fit" to be even larger.

4. Individual Value

The above examples illustrate that team performance is a function of both individual ability and teammate fit. In a case like this where there is no defense, no psychology, and tactics have been assumed away (by assuming that teams know who to start serving with and because there is no defense), team performance really is just this combination of player ability and fit. Explicitly, the results of the frescoball teams are just described by the complex operations of equation A16-A19 on individual probabilities. But “fit” and “ability” are more macro-level descriptions used by sports analysts and more useful without the detailed workings of the game being spelled out, as is often the case for other sports. We attempt here to quantify “fit” and “ability” in this game where tactics have been eliminated.

As mentioned above, it is commonly assumed that team performance is the linear sum of player ability. In this work where we wish to look at the difference between the whole and the sum of the parts, we then take the following relationship to be operating at a macro-level:

$$\text{Team Performance} = \text{Ability} + \text{Fit} \tag{7}$$

With the average scores in the above examples serving as Team Performance, we then only need to parse out Ability from Fit (we sometimes label “Ability” as “player value”). With fixed and known probabilities for all player types, Ability is fully determined, just not quantified into a single number as is usually done and as necessary for some direct solution to eq. 7. If these probabilities are reduced to a single number, then we have a known measure of Ability and eq. 7 can be solved for a magnitude of Fit, giving a sense of the relative magnitude of fit to ability.

The problem is that such a reduction of multiple skills to a single number is not obvious and a unique reduction may not be possible given the complex operations of eq. A16-A19. To reduce player ability to a single number, we simply use a linear regression approach, assuming that all of team performance is due to ability. This is an optimal way to model a team as the sum of its parts. Specifically, we propose to apply ordinary least squares regression on all 10 teams in a league against the number of players of each type on the team,

$$\text{Average Team Hits} = \alpha_{AC} N_{AC} + \alpha_{AI} N_{AI} + \alpha_{NC} N_{NC} + \alpha_{NI} N_{NI} + \varepsilon \tag{8}$$

where N represents the number of players on the team of the type shown in the subscript, ε is a measure of misfit of the linear model, and the coefficients α become the marginal hit values of each player.

DOI: 10.2202/1559-0410.1098

This is a specific instance of "opportunity value methods," (OVM) because it statistically estimates the value of a player *per opportunity* – however you choose to measure opportunity (seasons, minutes, games, plays, etc.) – by looking at team results during that opportunity. Our specific approach has distinct precedents in the practicing side of sports analysis. Rosenbaum (2004), Ilardi (2007), and WINVAL (e.g., Corazza, 2006) follow this exact principle in their works on basketball, applying regression to team results over segments of games where lineups are constant. Wolfers (2007) has suggested this approach on a more general level for identifying the range of player values in different sports.

The residuals of the regression (8) look like random variability to a statistical regression, but because of the known and fixed characteristics of the players in our Markov model, these residuals are simply the inability of the linear model to capture ability as a single number and, for our purposes, the residuals become a measure of teammate Fit. We call this a "naïve" linear regression because it doesn't know how frescoball works, in this case through the Markov model, and looks to estimate player value only from the results. Further, note that, because there is no real random variability in the Markov model results in the above examples, this is also an exercise in model "mis-fit" more than regression with typical real-world data. (Ironically, the "mis-fit" of the model reflects "fit" of teammates.) Whereas application of OVMs to empirical data yields residuals that reflect variability as well as the inappropriateness of the model for the game structure, application here provides residuals that reflect *only* the inappropriateness of linear model and, hence, the fit. We view this as a benefit of using regression, not a detriment in spite of regression often being applied to data with random variability.

Table 5
Regression-Estimated Values for Different Player Types

	Example 1	Example 2	Example 3
AC	3.90	6.56	4.22
AI	2.56	2.17	2.02
NC	2.69	3.92	3.72
NI	1.69	1.29	1.61

Using the 10-team results of example 1 and ordinary least-squares eq. 8[2], we find the coefficients representing player ability (per presence on a team) shown in Table 5. All 3 examples show AC as the best player and NI the worst, which was the intended setup of the examples. Further, in example 1, where AI

[2] The intercept is set to 0 in the regression. The results then have just the players and the residual, which we are calling fit, with no constant to interpret.

and NC fit better with different players, AI and NC are very comparable in ability using this technique. In examples 2 and 3, teams with NC always outperformed the comparable team with AI in its place and, as a result, Table 5 shows NC as a clearly better player.

What's then interesting from the regression are the residuals, which give us a measure of teammate fit relative to this measure of player ability. These are shown in Table 6, with the team score shown first and the residual shown in parentheses as a percentage of the team score.

Example 1 has results that are again reflective of what was noted above without regression. NC "fits" better with AC and NC – hence, the fit shown in parentheses for AC/NC and NC/NC is positive at 0.5% and 1.9%, respectively. AI "fits" better with AI and NI – hence, the entries for AI/AI and AI/NI show positive entries in parentheses at 2.4% and 2.8%, respectively. In this example, fit appears to be a small fraction of team performance, with no team showing more than 8% of their team performance as fit (AC/NI are 8.0% worse than they should be based on the calculated player values).

This is not the case in Example 2, where the fit percentages shown in Table 7 reflect much larger importance of fit. Several teams show fit percentages with an absolute magnitude over 10% and one team reaching 35%. In the original discussion of this example, we showed how the marginal value of AC is small when replacing one NI in an NI/NI team, but larger when replacing NI in an AC/NI team. This would imply a larger value of fit in AC/AC than in AC/NI and the regression results reflect this, where AC/AC has a fit of 11.4% and AC/NI has a fit of -35.5%. So this regression technique is capturing the story made above without regression.

Table 6
Team Average Scores for Example 1 with Residual (or Fit) in Parentheses, as a Percentage of Average Score

	AC	AI	NC	NI
AC	8.1 (3.6%)	6.3 (-3.1%)	6.6 (0.5%)	5.2 (-8.0%)
AI	6.3 (-3.1%)	5.2 (2.4%)	5.1 (-3.6%)	4.4 (2.8%)
NC	6.6 (0.5%)	5.1 (-3.6%)	5.5 (1.9%)	4.3 (-1.4%)
NI	5.2 (-8.0%)	4.4 (2.8%)	4.3 (-1.4%)	3.5 (4.9%)

Table 7
Team Average Scores for Example 2 with Residual (or Fit) in Parentheses, as a Percentage of Average Score

	AC	AI	NC	NI
AC	14.8 (11.4%)	7.0 (-25.0%)	10.9 (3.9%)	5.8 (-35.5%)
AI	7.0 (-25.0%)	5.0 (13.3%)	5.7 (-6.9%)	4.3 (19.0%)
NC	10.9 (3.9%)	5.7 (-6.9%)	8.1 (2.9%)	4.7 (-10.6%)
NI	5.8 (-35.5%)	4.3 (19.0%)	4.7 (-10.6%)	3.5 (25.2%)

Table 8
Team Average Scores for Example 3 with Residual (or Fit) in Parentheses, as a Percentage of Average Score

	AC	AI	NC	NI
AC	8.9 (5.6%)	6.0 (-4.4%)	7.8 (-1.6%)	5.2 (-11.9%)
AI	6.0 (-4.4%)	4.2 (4.6%)	5.3 (-8.7%)	4.0 (8.4%)
NC	7.8 (-1.6%)	5.3 (-8.7%)	7.9 (6.3%)	4.9 (-8.5%)
NI	5.2 (-11.9%)	4.0 (8.4%)	4.9 (-8.5%)	3.6 (9.8%)

Finally, Table 8 shows the results of example 3 with the associated fit percentages. This example had the odd result that NC/NC was better than NC/AC. Even though Table 5 still shows that AC is the best player in this group, the manufactured characteristics of NC show that a team of NC/NC gets a 6.3% benefit of fit, boosting that team to a higher level than the AC/NC team with a -1.6% fit. This only works because the apparent values of AC and NC, as evaluated by the regression, are close: 4.22 to 3.72. So an edge in fit between NC/NC over NC/AC overcomes the 0.5 hits advantage in ability of AC over NC, with the absolute magnitude of fit in AC/NC being -0.1 and in NC/NC being 0.5.

5. Individual Value with League Context

As an additional exercise, we now extend the regression technique outlined above to look at larger synthesized leagues of players, generate results, and estimate player value. The concept of players rotating through all possible teammates to better determine player value is a concept suggested by Oliver (Chp 15, 2004), but cannot be achieved in real leagues. As a result, we will generate a few "seasons" of each league with each season involving some player movement. This limits the number of contexts available to estimate player value and fit using regression, but is more realistic for sporting leagues where players do not rotate through all teams. We wish to understand how much this limitation affects *estimates* of player value and fit.

Specifically, we develop two leagues of 50 players each, which yields 25 teams in each league. The players' transition probabilities are randomly generated with the constraints of eqs. 5 and 6, and given the following:

$$0.9 \leq p_{11} + p_{12} < 0.99$$

$$0.5 \leq p_{21} + p_{22} < 0.9$$

We then simulate 8 "seasons" of the 25 teams (with no mid-season trades), where more successful teams are simulated to be more stable from season to season than less successful teams by assigning a lower probability of swapping a teammate (which teammate to be swapped was randomly assigned). In each league, we assign a maximum probability of staying with the team for the next season of 95%. The actual probability of swapping is set to be linear with respect to how much worse a team's average score was relative to the best and worst (of all teams simulated through all league seasons). With a total of 1225 possible teams in each league, these 8 seasons generate 200 total teams, of which about 80% and 75% in League 1 and League 2, respectively, end up unique with the player movement rules. This is then sampling approximately 10-15% of possible teams in a league, a fairly large percentage compared to what you would see in professional teams of multiple players.

In each of the two leagues, we then apply the value/fit estimation procedure of Section 4 both on the sample of 8 seasons and on the full 1225 possible teams. Differences in the results between the 8-season estimates and the full estimates reflect how the OVM technique can be fooled by not seeing a full array of player movement.

Looking first at League 1, Table 9 shows the probabilities used for all players as well as the estimated value for the players using both the full league ("full" in Table 9) and the 8 season sub-sample ("8 seas"). In this league, 41 of 50 players are estimated to be within one standard error of the full league estimate and 46 of 50 are within two standard errors. These are not highly discrepant with theory and, hence, the methodology is not dramatically misled by the limited player movement of this league.

Table 9
Player Parameters and Value Estimates, League 1

Player	p_{11}	p_{12}	p_{21}	p_{22}	Estd. Value, hits (full)	Estd. Value, hits (8 seas)	SE
1	58%	38%	29%	42%	3.63	3.66	0.65
2	1%	91%	0%	81%	2.12	1.98	0.66
3	54%	40%	8%	68%	3.23	3.26	0.66
4	45%	46%	36%	41%	3.26	3.49	0.64
5	53%	38%	9%	69%	2.97	2.33	0.67
6	46%	51%	21%	45%	2.83	3.01	0.65
7	59%	34%	17%	37%	2.30	2.19	0.64
8	44%	54%	27%	27%	2.46	2.35	0.64
9	27%	71%	26%	54%	3.34	3.11	0.65
10	32%	59%	30%	22%	1.58	1.55	0.65
11	25%	67%	4%	65%	1.92	1.51	0.65
12	86%	8%	33%	25%	4.27	3.88	0.68
13	63%	31%	41%	12%	2.94	3.64	0.71
14	92%	6%	35%	20%	5.49	4.89	0.69
15	22%	69%	13%	56%	1.93	1.72	0.66
16	55%	35%	54%	10%	2.93	2.62	0.67
17	83%	11%	67%	17%	7.24	7.05	0.66
18	34%	63%	18%	72%	3.99	4.87	0.66
19	81%	12%	54%	18%	5.24	4.60	0.71
20	23%	71%	15%	55%	2.18	3.34	0.70
21	58%	36%	36%	23%	2.93	3.07	0.66
22	19%	79%	16%	53%	2.34	2.73	0.65
23	96%	2%	7%	52%	5.20	17.30	1.32
24	38%	53%	12%	70%	2.83	0.69	0.70
25	5%	86%	2%	55%	1.11	1.59	0.65
26	28%	66%	12%	54%	2.03	2.28	0.67
27	10%	85%	8%	56%	1.72	1.77	0.77
28	54%	40%	49%	39%	5.23	5.40	0.76
29	38%	61%	28%	31%	2.53	3.06	0.65
30	52%	43%	44%	11%	2.65	2.49	0.66
31	76%	21%	50%	3%	3.82	4.51	0.68
32	33%	59%	5%	54%	1.62	2.12	0.64
33	80%	15%	33%	48%	5.37	5.47	0.67
34	67%	31%	42%	41%	5.71	6.66	0.68
35	80%	11%	70%	19%	7.25	7.90	0.67
36	17%	77%	16%	34%	1.33	1.39	0.64
37	11%	84%	0%	66%	1.66	1.59	0.66
38	28%	69%	8%	76%	3.14	3.20	0.66
39	71%	23%	31%	32%	3.48	2.96	0.65
40	59%	38%	24%	32%	2.90	2.72	0.65

Table 9
Player Parameters and Value Estimates, League 1

Player	p_{11}	p_{12}	p_{21}	p_{22}	Estd. Value, hits (full)	Estd. Value, hits (8 seas)	SE
41	98%	0%	25%	30%	5.99	18.63	1.32
42	44%	54%	25%	28%	2.33	0.27	0.68
43	53%	41%	27%	39%	2.80	2.60	0.67
44	8%	87%	0%	58%	1.38	1.96	0.64
45	75%	19%	18%	44%	3.36	3.35	0.69
46	15%	76%	15%	36%	1.15	1.17	0.64
47	83%	10%	33%	18%	3.38	3.93	0.70
48	80%	14%	56%	16%	5.31	5.08	0.64
49	36%	55%	34%	24%	2.03	3.20	0.66
50	57%	37%	5%	62%	2.79	2.65	0.66

The biggest errors observed between player values estimated from the sub-league are with players 23 and 41 who were randomly paired together in the 8 season sub-league in the second season and are a perfect fit for one another, hitting very consistently back and forth to each other's sweet spot, despite not being very athletic in the way we've defined it. The 37.7 score of this team dwarfs that of other teams and can make each of them look particularly good. The "fit" of the two players based on the full league valuation is 26.5 (= 37.7 – 5.2 – 6.0 = team score – marginal value of player 1 – marginal value of player 2). But using just the 8 season sub-league, which doesn't have the benefit of all perspectives, the players look a lot better than they are, with a "fit" of just 1.7 (= 37.7 – 17.3 – 18.6). Acquiring player 23 or 41 (and paying a lot of money) because they appear to be so good in the 8 season sub-league could be a very significant error for other teams. Their respective first seasons with other teammates show this clearly. The team of player 23 and player 24 scored only 5.9 expected hits, and the team of player 41 and player 42 scored only 6.8 in that first season. Both players, however, do rank highly by the full method, player 23 ranking 10th and player 41 ranking 3rd.

Because League 1 randomly drew the best possible fit combination in its 8 season sub-sample, we create a second League to try to avoid this. League 2 has the same general constraints as League 1, but with different random draws of player probabilities and combinations. This realization yields a maximum team score over the 8 seasons of 17.9, well below the maximum over all combinations of 35.9. Table 10 shows the probabilities and value estimates for the players in this league. In this league, only 31 of 50 players are estimated to be within one standard error of the full league estimate and 45 of 50 are within two standard errors. These totals are slightly less than would be expected statistically. However, the correlation between the estimates from the full league and the 8 seasons of data is high at 95%, implying that the order is reasonably maintained even if the absolute value estimates are off. This was not the case in League 1, where correlation was only 69% because of that unique fit situation.

Table 10
Player Parameters and Value Estimates, League 2

Player	p11	p12	p21	p22	Estd. Value, hits (full)	Estd. Value, hits (8 seas)	SE
1	37%	61%	7%	53%	2.12	2.15	0.48
2	54%	37%	0.2%	62%	1.97	1.87	0.50
3	85%	13%	15%	64%	6.21	7.92	0.51
4	85%	13%	57%	17%	7.18	7.88	0.51
5	47%	50%	40%	23%	2.96	3.00	0.47
6	63%	30%	42%	9%	2.60	3.21	0.49
7	9%	85%	6%	74%	2.30	2.00	0.50
8	88%	7%	32%	47%	6.24	8.06	0.52
9	90%	5%	72%	11%	8.71	8.18	0.53
10	29%	63%	25%	28%	1.49	1.57	0.47
11	65%	26%	25%	28%	2.32	3.14	0.55
12	76%	15%	42%	26%	3.85	3.09	0.56
13	63%	31%	38%	29%	3.38	3.24	0.49
14	24%	68%	16%	58%	2.24	2.45	0.47
15	18%	78%	5%	79%	2.83	2.50	0.46
16	52%	44%	38%	20%	2.78	2.78	0.48
17	92%	4%	9%	49%	4.49	4.90	0.51
18	40%	55%	36%	23%	2.33	1.99	0.50
19	13%	86%	5%	52%	1.44	1.10	0.48
20	58%	38%	4%	81%	4.24	3.81	0.47
21	34%	57%	30%	58%	3.81	4.40	0.47
22	25%	73%	3%	49%	1.47	1.79	0.48
23	30%	64%	6%	44%	1.31	1.42	0.46
24	18%	75%	12%	40%	1.14	0.83	0.47
25	70%	24%	59%	11%	4.50	4.02	0.48
26	33%	60%	16%	72%	3.59	4.67	0.52
27	96%	2%	16%	41%	5.44	5.94	0.50
28	12%	82%	11%	42%	1.19	1.19	0.47
29	51%	41%	1%	89%	3.82	4.34	0.46
30	19%	72%	12%	45%	1.22	1.78	0.46
31	25%	71%	9%	70%	2.70	3.09	0.49
32	36%	54%	25%	46%	2.27	2.53	0.50
33	79%	11%	3%	59%	2.87	2.38	0.51
34	44%	52%	12%	41%	1.97	1.94	0.48
35	82%	14%	50%	9%	4.77	3.90	0.47
36	12%	87%	3%	80%	2.89	3.32	0.47

Table 10
Player Parameters and Value Estimates, League 2

Player	p11	p12	p21	p22	Estd. Value, hits (full)	Estd. Value, hits (8 seas)	SE
37	44%	48%	1%	63%	2.00	1.61	0.54
38	54%	41%	9%	81%	4.51	4.85	0.53
39	59%	37%	47%	23%	3.91	3.66	0.48
40	40%	51%	20%	35%	1.65	1.58	0.46
41	56%	40%	39%	39%	4.28	3.54	0.50
42	17%	76%	17%	68%	3.00	3.40	0.48
43	64%	31%	22%	30%	2.53	3.28	0.51
44	95%	4%	64%	18%	10.96	9.08	0.52
45	28%	69%	17%	61%	3.03	3.61	0.48
46	24%	72%	21%	68%	3.82	3.58	0.50
47	74%	23%	7%	69%	4.31	3.97	0.52
48	88%	10%	71%	4%	8.57	11.19	0.55
49	69%	22%	19%	44%	2.74	3.32	0.49
50	18%	80%	4%	73%	2.56	2.41	0.47

In both Leagues, the average player value is estimated too high from the 8 season data. In League 1, the average player was estimated as being worth 3.74 hits from the 8 season data, but estimated as 3.22 hits from the full league data. In League 2, it was 3.63 vs 3.49. This is very sensible as teams with success – whether due to talent or fit tend to stay together longer, thus blurring the reasons for the success. This generally implies that player contribution to team success may get overestimated when player movement is less on teams that are good.

Further emphasizing this, we calculate the average fit of teams using the 8-season data and the full set of possible teams for both leagues. In League 1, the average fit is estimated as 3.4% using full data. If one were to use just the 8 season data, the average fit would be 0% because of regression yielding mean 0 residuals. In League 2, we see the same pattern, with a 2.3% fit average from the full data. The player movement rules in the 8-season data lead to a bias towards good fitting teammates, but our regression to estimate player value assumes 0 fit over the subset of teams and, hence, overestimates player value.

There are two other characterizations of the impact of fit worth mentioning. First, let's look at how many teams each team would beat with and without fit. For instance, a team having players 1 and 2 in League 1 has an average score of 5.73 in reality, but the "sum of players" leads to an average score of 5.75. In League 1, that team beats 587 out of 1224 other teams (0.480), but comparing the sum of players to other sums of players, the team beats 540 of 1224 teams (0.441), for a deviation of about 4%. If we look across all potential combinations (the full league), we see the average absolute change in this surrogate winning percentage is 6.9%. In League 2, this number is 7.2%. Though that number may not apply to other sports, 7% of a full season in other sports is a pretty sizable impact. The second related characterization of fit is looking at how the "sum of players" would do in the league (not against the other "sums of players"). The real team wins minus the "sum of player" wins tells us how much fit impacts that team. As above, the average absolute impact is about 7-8% (7% in League 1 and 8% in League 2) on this surrogate winning percentage. In Tables 11 and 12, we break things down a little more, however. We look at the average absolute deviation in this surrogate win% across the teams that each player is on. So the league average deviation is about 7%, but Player 12 in League 2 sees about a 4.5% average absolute win percentage change across his teams, meaning that he is less context sensitive than Player 46, who sees about a 12% average absolute win percentage change. These two players are almost identical in terms of estimated value, but there is less variation in what the team does with Player 12 than with Player 46. Though we don't carry out analysis of why certain players have more context-dependence, we present the underlying probabilities for all players in the table for the reader's inspection. By inspection, it does appear that extreme players in hitting shots from Zone 1 are the ones most prone to context-dependence.

Table 11
Player Parameters and Average Absolute Fit Impact, League 1

Player	p11	p12	p21	p22	Avg. Absolute Impact of Fit	Estd. Value, hits (full)
1	58%	38%	29%	42%	2.1%	3.63
2	1%	91%	0%	81%	16.3%	2.12
3	54%	40%	8%	68%	4.1%	3.23
4	45%	46%	36%	41%	7.7%	3.26
5	53%	38%	9%	69%	5.5%	2.97
6	46%	51%	21%	45%	3.7%	2.83
7	59%	34%	17%	37%	4.4%	2.30
8	44%	54%	27%	27%	4.1%	2.46
9	27%	71%	26%	54%	9.2%	3.34
10	32%	59%	30%	22%	7.0%	1.58
11	25%	67%	4%	65%	9.2%	1.92
12	86%	8%	33%	25%	11.7%	4.27
13	63%	31%	41%	12%	4.7%	2.94
14	92%	6%	35%	20%	17.4%	5.49
15	22%	69%	13%	56%	10.5%	1.93
16	55%	35%	54%	10%	6.8%	2.93
17	83%	11%	67%	17%	2.6%	7.24
18	34%	63%	18%	72%	10.4%	3.99
19	81%	12%	54%	18%	5.3%	5.24

Table 11
Player Parameters and Average Absolute Fit Impact, League 1

Player	p11	p12	p21	p22	Avg. Absolute Impact of Fit	Estd. Value, hits (full)
20	23%	71%	15%	55%	9.8%	2.18
21	58%	36%	36%	23%	3.2%	2.93
22	19%	79%	16%	53%	9.2%	2.34
23	96%	2%	7%	52%	20.1%	5.20
24	38%	53%	12%	70%	10.8%	2.83
25	5%	86%	2%	55%	8.7%	1.11
26	28%	66%	12%	54%	8.0%	2.03
27	10%	85%	8%	56%	9.5%	1.72
28	54%	40%	49%	39%	5.5%	5.23
29	38%	61%	28%	31%	4.4%	2.53
30	52%	43%	44%	11%	4.6%	2.65
31	76%	21%	50%	3%	8.9%	3.82
32	33%	59%	5%	54%	6.6%	1.62
33	80%	15%	33%	48%	4.6%	5.37
34	67%	31%	42%	41%	4.2%	5.71
35	80%	11%	70%	19%	2.5%	7.25
36	17%	77%	16%	34%	7.4%	1.33
37	11%	84%	0%	66%	9.9%	1.66
38	28%	69%	8%	76%	10.5%	3.14
39	71%	23%	31%	32%	3.7%	3.48
40	59%	38%	24%	32%	4.0%	2.90
41	98%	0%	25%	30%	20.8%	5.99
42	44%	54%	25%	28%	4.3%	2.33
43	53%	41%	27%	39%	3.4%	2.80
44	8%	87%	0%	58%	8.6%	1.38
45	75%	19%	18%	44%	6.0%	3.36
46	15%	76%	15%	36%	7.8%	1.15
47	83%	10%	33%	18%	10.1%	3.38
48	80%	14%	56%	16%	4.9%	5.31
49	36%	55%	34%	24%	6.9%	2.03
50	57%	37%	5%	62%	2.8%	2.79

One thing we wanted to do with these leagues is try to suggest ideas for finding fit in real data. We look at the top 10 scoring teams in each league using the 8 season data and their fit value from the full data set. Because we won't know actual fit values for real teams, but we could see the best teams in a real league, would those teams have higher values of fit? If so, we could focus a study on the best teams in a real league to extract real values of fit. In fact, this is the case in our synthetic leagues. Table 13 shows the top 10 unique teams in each league and the magnitude of fit on those teams. In League 1, the top team clearly has a significant amount of fit and, with only little inspection, would beg for an investigation of fit. The other teams in the Top 10 are not nearly as high, but still elevated. A one-tail t-test of fit values between the remainder of the top 10 and those teams outside the top 10 in League 1 is significant at a 99% level. In League 2, without a striking outlier, a similar t-test does suggest that fit values in the top 10 teams are significantly higher than the fit values for teams below them. Both leagues then have indications that those top teams have higher fit, suggesting where to look for it in real data.

Table 12
Player Parameters and Average Absolute Fit Impact, League 2

Player	p11	p12	p21	p22	Avg. Absolute Impact of Fit	Estd. Value, hits (full)
1	37%	61%	7%	53%	6.1%	2.12
2	54%	37%	0.2%	62%	6.3%	1.97
3	85%	13%	15%	64%	9.6%	6.21
4	85%	13%	57%	17%	7.1%	7.18
5	47%	50%	40%	23%	4.4%	2.96
6	63%	30%	42%	9%	8.8%	2.60
7	9%	85%	6%	74%	13.5%	2.30
8	88%	7%	32%	47%	6.9%	6.24
9	90%	5%	72%	11%	3.2%	8.71
10	29%	63%	25%	28%	7.5%	1.49
11	65%	26%	25%	28%	7.9%	2.32
12	76%	15%	42%	26%	4.5%	3.85
13	63%	31%	38%	29%	3.3%	3.38
14	24%	68%	16%	58%	11.5%	2.24
15	18%	78%	5%	79%	12.7%	2.83
16	52%	44%	38%	20%	5.6%	2.78
17	92%	4%	9%	49%	16.1%	4.49
18	40%	55%	36%	23%	6.2%	2.33
19	13%	86%	5%	52%	8.5%	1.44
20	58%	38%	4%	81%	8.2%	4.24
21	34%	57%	30%	58%	11.3%	3.81
22	25%	73%	3%	49%	7.4%	1.47
23	30%	64%	6%	44%	7.2%	1.31
24	18%	75%	12%	40%	7.8%	1.14
25	70%	24%	59%	11%	3.0%	4.50
26	33%	60%	16%	72%	10.8%	3.59
27	96%	2%	16%	41%	20.0%	5.44
28	12%	82%	11%	42%	8.4%	1.19

Table 12
Player Parameters and Average Absolute Fit Impact, League 2

Player	p11	p12	p21	p22	Avg. Absolute Impact of Fit	Estd. Value, hits (full)
29	51%	41%	1%	89%	9.9%	3.82
30	19%	72%	12%	45%	8.5%	1.22
31	25%	71%	9%	70%	10.3%	2.70
32	36%	54%	25%	46%	9.9%	2.27
33	79%	11%	3%	59%	6.9%	2.87
34	44%	52%	12%	41%	6.2%	1.97
35	82%	14%	50%	9%	10.3%	4.77
36	12%	87%	3%	80%	11.9%	2.89
37	44%	48%	1%	63%	6.8%	2.00
38	54%	41%	9%	81%	9.8%	4.51
39	59%	37%	47%	23%	1.6%	3.91
40	40%	51%	20%	35%	7.2%	1.65
41	56%	40%	39%	39%	3.1%	4.28
42	17%	76%	17%	68%	14.1%	3.00
43	64%	31%	22%	30%	7.4%	2.53
44	95%	4%	64%	18%	4.8%	10.96
45	28%	69%	17%	61%	8.7%	3.03
46	24%	72%	21%	68%	11.6%	3.82
47	74%	23%	7%	69%	5.5%	4.31
48	88%	10%	71%	4%	6.2%	8.57
49	69%	22%	19%	44%	5.9%	2.74
50	18%	80%	4%	73%	9.9%	2.56

Other methods of player valuation besides regression could, of course, be used in this analysis (and the analysis of the 3 examples). Those other methods would lead to different absolute results. Our goal here was to use a fairly simple and accurate model of condensing player value to a single number; in this case, ordinary least squares regression provides an optimal estimate given the linear model of team production. Player evaluation methods that use the underlying transition probabilities (this would not be "naïve", as we desired above) would lead to no difference in estimates of value/fit between the full league and the 8 season data. This requires full knowledge and characterization of skills (the probabilities in frescoball); this may be possible in frescoball, but not typically in other sports. Further, in cooperative events, like hitting the ball back and forth to each other, like running the ball over left tackle, like defending a pick-and-roll, it's often hard to see who is to blame or credit in a play's failure or success. As such, personnel evaluators do tend to look at team results alone to try to parse the information, which is the basic nature of our simple value estimation method.

Table 13
Team Average Scores and Fits for Top 10 Teams in Each of Simulated Leagues

League	Player 1	Player 2	Avg Score	Fit Value
1	23	41	37.7	26.46
1	18	35	13.1	1.88
1	33	34	12.8	1.67
1	17	48	12.4	-0.18
1	17	18	11.8	0.62
1	14	31	11.8	2.52
1	18	34	11.4	1.69
1	17	45	10.8	0.21
1	12	48	10.7	1.14
1	35	42	10.6	0.98
2	3	4	17.9	4.50
2	17	48	17.5	4.46
2	8	27	17.3	5.59
2	47	48	15.2	2.32
2	43	44	13.2	-0.24
2	3	35	13.1	2.16
2	4	47	13.0	1.49
2	6	44	12.8	-0.78
2	41	48	12.5	-0.31
2	9	43	11.6	0.32

6. Conclusions

Though casual or pro-hopeful frescoball players may view the above results to suggest that consistency is more important than athleticism to be common across partners to enhance teammate fit, our goal was not to particularly evaluate frescoball. The parameterizations of the model are not based on data from actual frescoball players, as we could not find data.

Rather, the principal thrust of this work is to illustrate that fit does affect team performance and to quantify it even in a synthetic realm. Further, we wish to frame the discussion of fit so that future work can be done using more complex games and parameters taken from documented statistics. Specifically, we wish to emphasize the following definitions and points made above:

- Fit can be defined as the difference between the sum of the parts and the whole. If the whole is greater than the sum of the parts, then fit is positive. In sports, the "whole" is the productivity of the team in terms of wins or, in this paper, the expected number of hits of the team.

Likewise, the "parts" are the player abilities, player ratings, or player productivity, however it is phrased. Fit can also be described as the complementarity of skills and is a function of how a team is constructed.

- One characteristic of teammate fit is that teams can trade players and both teams see improvement (or decline). Though this may be attributed to "chemistry" (or psychology) in the real world, we showed how it can occur simply because of skills matching in the synthesized Markov model of frescoball and this likely is part of the explanation of such occurrences in more complicated sports.
- A more basic characteristic of fit is that a player's marginal value to his team is context dependent. This was illustrated with and without a regression technique to estimate player value. Studies assuming a context-independent marginal value of a player (common in sports economics, e.g., Scully, 1974, Krautmann, 1999) are using a simplification whose accuracy can be questioned. In particular, our study suggests that a player's marginal economic value to a team, which relates to salary, is a function of their own skill, the size of the market they play in, other economic factors (e.g., star power, as shown having limited impact in Berri and Schmidt, 2004), *and* the teammates they will have.
- When there is little player movement, player value estimates can be influenced by context. Our league studies showed that player values estimated through regression led to sometimes misleading (and often optimistic) numbers when using just a subset of possible teams sampled. Similarly, if there is a bias towards keeping successful teams together and separating bad teams, this is equivalent to introducing a bias for positive fitting teams. Hence, the average fit across teams is probably positive and falsely credited to player ability if the player evaluation method assumes fit is 0. Though the process of evaluating player ability is more complex than this, we suppose that context can influence many perceptions, as described in Gladwell (2002, p. 160) as "human beings invariably make the mistake of overestimating the importance of fundamental character traits and underestimating the importance of the situation and context," and in general by numerous peer-reviewed studies (e.g., Tversky and Kahneman, 1974.).
- There are certain synthetic players who appear to be more subject to context than others. Those players were not studied thoroughly here, but appear to be more extreme in their skills or lack thereof.
- With the synthetic frescoball examples and simulated leagues, fit can be a very significant component of team performance. The best teams

tended to have the largest amount of fit. In attempting to find evidence of and quantify fit with real data, we would suggest looking at the best teams.

It has not been demonstrated that the above conclusions are relevant in other sports, but the mere presence of "fit" in this game, with general skills that we can reasonably label as "athleticism" and "consistency," suggests relevance in other sports.

At a minimum, we hope that we framed the discussion of fit and provided a strategy to look for it in other sports.

Appendix A: Derivation of the Fundamental Matrix Solution

We start with the matrix, (**I** – **Q**).

$$\mathbf{B} = \mathbf{I} - \mathbf{Q} = \begin{bmatrix} 1 & 0 & -p_{113} & -p_{114} \\ 0 & 1 & -p_{123} & -p_{124} \\ -p_{211} & -p_{212} & 1 & 0 \\ -p_{221} & -p_{222} & 0 & 1 \end{bmatrix} \tag{A1}$$

To obtain a closed-form analytical solution to this matrix, we can subdivide **B** into a partitioned matrix made up of four submatrices

$$\mathbf{B} = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} \tag{A2}$$

where

$$B_{11} = B_{22} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \tag{A3}$$

$$B_{12} = \begin{bmatrix} -p_{113} & -p_{114} \\ -p_{123} & -p_{124} \end{bmatrix} \tag{A4}$$

$$B_{21} = \begin{bmatrix} -p_{211} & -p_{212} \\ -p_{221} & -p_{222} \end{bmatrix} \tag{A5}$$

This allows us to employ a matrix identity for the Fundamental Matrix, **F**,

$$\mathbf{B}^{-1} = \mathbf{F} \tag{A6}$$

from Schweppe (1973) which states

$$\begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}^{-1} = \begin{bmatrix} F_{11} & F_{12} \\ F_{21} & F_{22} \end{bmatrix} \tag{A7}$$

where

$$F_{11} = B_{11}^{-1} + B_{11}^{-1} B_{12} \left[B_{22} - B_{21} B_{11}^{-1} B_{12} \right]^{-1} B_{21} B_{11}^{-1} \tag{A8}$$
$$F_{12} = -B_{11}^{-1} B_{12} \left[B_{22} - B_{21} B_{11}^{-1} B_{12} \right]^{-1} \tag{A9}$$
$$F_{21} = -\left[B_{22} - B_{21} B_{11}^{-1} B_{12} \right]^{-1} B_{21} B_{11}^{-1} \tag{A10}$$
$$F_{22} = \left[B_{22} - B_{21} B_{11}^{-1} B_{12} \right]^{-1} \tag{A11}$$

On the surface, this may appear to have gained little. However, we have reduced the 4 x 4 matrix into a series of 2 x 2 matrices. A direct closed-form analytical solution for a 4 x 4 matrix is complicated, whereas each individual 2 x 2 matrix can be inverted very easily. Furthermore, note the common term (an inverted term in square brackets) in each of equations A8 through A11 and defined as

$$\Xi = \left[B_{22} - B_{21} B_{11}^{-1} B_{12} \right]^{-1} \tag{A12}$$

In closed form, we can calculate Ξ as the 2 x 2 matrix

$$\begin{bmatrix} \frac{-1 + p_{221}p_{114} + p_{222}p_{124}}{\zeta} & -\frac{p_{211}p_{114} + p_{212}p_{124}}{\zeta} \\ -\frac{p_{221}p_{113} + p_{222}p_{123}}{\zeta} & \frac{-1 + p_{211}p_{113} + p_{212}p_{123}}{\zeta} \end{bmatrix} \tag{A13}$$

where

$$\begin{aligned} \zeta = -1 &+ p_{221}p_{114} + p_{222}p_{124} + p_{211}p_{113} - p_{211}p_{113}p_{222}p_{124} \\ &+ p_{212}p_{123} - p_{212}p_{123}p_{221}p_{114} + p_{211}p_{114}p_{222}p_{123} \\ &+ p_{212}p_{124}p_{221}p_{113} \end{aligned} \tag{A14}$$

Now, also noting that

$$B_{11}^{-1} = B_{22}^{-1} = B_{11} = B_{22} = \mathbf{I}_2 \tag{A15}$$

DOI: 10.2202/1559-0410.1098

where $\mathbf{I}_2$ is a 2 x 2 identity matrix, we can restate Eqs. A8 through A11 as

$$F_{11} = B_{11}^{-1} + B_{12}\Xi B_{21} \tag{A16}$$
$$F_{12} = -B_{12}\Xi \tag{A17}$$
$$F_{21} = -\Xi B_{21} \tag{A18}$$
$$F_{22} = \Xi \tag{A19}$$

In summary, the solution for the terms of this matrix proceeds as follows:

1. Calculate Ξ using Eqs. A13 and A14.
2. Use Ξ in equations A16 through A19 to calculate the terms of **F** as partitioned in Eq. A7.
3. By virtue of Eq. A6, the solution for the matrix is complete.
4. Once assembled, the resulting matrix **F** is 4 x 4 and its rows may be summed or other relevant operations performed.

References

Alamar, Ben, Ma, Jeff, Desjardins, Gabriel M., and Ruprecht, Lucas. "Who Controls the Plate? Isolating the Pitcher/Batter Subgame." *J. Quantitative Analysis in Sports*, 2006, vol 2(3), article #4.

Berri, David J., Schmidt, Martin B., and Brook, Stacey L. "Stars at the Gate: The Impact of Star Power on NBA Gate Revenues." 2004. *Journal of Sports Economics*, vol 5 (1), 33-50.

Corazza, Ryan. "Jeff Sagarin: the human sports rating system" Indiana Student News. 2/2/2006. http://www.idsnews.com/news/story.aspx?id=33683

Crossman, Matt. "Life Without Vick." *The Sporting News*. 2007. http://www.sportingnews.com/yourturn/viewtopic.php?t=244201.

Fry, Michael J., Lundberg, Andrew W., and Ohlmann, Jeffrey W. "A Player Selection Heuristic for a Sports League Draft." *J. Quantitative Analysis in Sports*, 2007, vol 3(2), article #5.

Gladwell, Malcolm. *The Tipping Point*. 2002. Back Bay Books, New York, NY.

Ilardi, Steve. "Adjusted Plus-Minus: An Idea Whose Time Has Come." 2007. *82games.com*, http://www.82games.com/ilardi1.htm.

James, Bill. *Win Shares*. 2002. STATS Publishing.

Krautmann, Anthony. "What's Wrong with Scully-Estimates of a Player's Marginal Revenue Product." Economic Inquiry, 1999, vol 37 (2), 369-381.

Noubary, Reza D. "Probabilistic Analysis of a Table Tennis Game." *J. Quantitative Analysis in Sports*, 2007, vol 3(1), article #4.

Oliver, Dean. *Basketball on Paper*. 2004. Brassey's, Washington, DC.

Page, Scott. *The Difference*. 2007. Princeton University Press, Princeton, NJ.

Rosenbaum, Dan T. "Measuring How NBA Players Help their Teams Win." 2004. *82games.com*, http://www.82games.com/comm30.htm.

Schatz, Aaron, Editor. *Pro Football Prospectus*. 2006. Workman Publishing Company, New York, NY.

Schweppe, F.C. "Uncertain Dynamic Systems." 1973. Prentice-Hall, Engelwood Cliffs, NJ.

Scully, Gerald. "Pay and Performance in Major League Baseball." *American Economic Review*, 1974, vol 64 (6), 915-930.

Totterdell, P. "Catching moods and hitting runs: mood linkage and subjective performance in professional sports teams." *Journal of Applied Psychology*, 2000, vol 85(6), 848-859.

Tversky, Amos and Kahneman, Daniel. "Judgment under uncertainty: Heuristics and biases." *Science*, 1974, vol 185, 1124-1131.

Wolfers, Stephen. Personal Communication. 2007.