

On the Development of a Soccer Player Performance Rating System for the English Premier League

Ian G. McHale

Centre for Operations Management, Management Science and Statistics, and Centre for Gambling Studies, Salford Business School, University of Salford, Salford M5 4WT, United Kingdom, i.mchale@salford.ac.uk

Philip A. Scarf

Centre for Operations Management, Management Science and Statistics, Salford Business School, University of Salford, Salford M5 4WT, United Kingdom, p.a.scarf@salford.ac.uk

David E. Folker

Football DataCo Ltd., London W1U 8PL, United Kingdom, dfolker@football-dataco.com

The EA Sports Player Performance Index is a rating system for soccer players used in the top two tiers of soccer in England—the Premier League and the Championship. Its development was a collaboration among professional soccer leagues, a news media association, and academia. In this paper, we describe the index and its construction. The novelty of the index lies in its attempts to rate all players using a single score, regardless of their playing specialty, based on player contributions to winning performances. As one might expect, players from leading teams lead the index, although surprises happen.

Key words: sports and recreation; football; performance measurement; ranking.

History: This paper was refereed. Published online in *Articles in Advance* April 6, 2012.

Performance assessment is a fundamental tool for quantitative analysts and operational researchers. For example, fund managers assess the performance of traders (Grinblatt and Titman 1989), and engineers monitor the performance of production lines (Swamidass and Newell 1987). In sports, performance measurement typically takes the form of a rating system (i.e., players receive points for performances) or a rankings list (i.e., players' performances are ordered). Many stakeholders in the sports industry, including the teams, fans, and pundits, use rating systems. In baseball, for example, the Oakland Athletics used a quantitative analyst to value players (Lewis 2003). The coaching staff used the values to identify and rank the most desirable players in the draft. By drafting and signing players who were effectively undervalued by the market, the team was able to compete with much wealthier teams for a fraction of the cost. Fans use official ratings to gauge how well their favourite player or team is performing, and then

debate rating and ranking lists among friends or on social networking websites.

Rating systems already exist in many sports. Individual sports (e.g., tennis) provide a simple framework for rating individual players. The official world rankings of the Association of Tennis Professionals (ATP) are derived from results in tournaments from the previous 52 weeks of competition. Its website (<http://www.atpworldtour.com/Rankings/Singles.aspx>) lists the most current rankings. The tournament's prestige and a player's progress through the tournament determines how points are awarded. Although the ATP world rankings are well established, they have been called into question in the academic literature. McHale and Morton (2011) show that the points that underlie these rankings perform poorly as a forecasting tool. Golf has a similar story; McHale and Forrest (2005) show the Official World Golf rankings list to be a relatively poor predictor of the result of the next tournament. It

appears, then, that ranking in individual sports is not simple.

Rating teams in team sports is similar to rating individuals in individual sports. It is often a simpler exercise because fewer teams are competing against each other, and one can organise and design leagues with the objective of ranking (and to some extent rating) the teams at the end of the season. For example, in a round-robin league, each team plays the other teams in turn and points are awarded for success. The National Football League regional leagues are examples of a round-robin league, and the final league table provides a list of ratings for the teams in that region. Where teams do not play in leagues, sports associations have typically commissioned consultants to produce ratings, which the sports associations publish; for example, the Fédération Internationale de Football Association (FIFA) publishes current FIFA ratings on its website (<http://www.fifa.com/worldfootball/ranking/index.html>). The International Cricket Council (ICC) publishes world rankings for national cricket teams for various formats of the game on its website (<http://www.cricinfo.com/rankings/content/current/page/211271.html>). Although these systems are referred to as ranking lists, they provide ratings that somewhat represent team strengths. Whether official rankings provide accurate estimates of team strength is an issue that has been addressed in various sports. The issue is not trivial, and the importance of such rating systems should not be understated. For example, the FIFA world rankings form part of the criteria for awarding work permits for players outside the European Union in the English Premier League (PL). McHale and Davies (2007) show that when it comes to forecasting match outcomes, the world rankings do not perform well. In other sports, tournament organisers use the official rankings for seeding (e.g., the US Open in tennis), so that the rankings can influence the tournament outcome. They might even use rankings to determine qualifications for a tournament; the top ten teams qualify automatically for the ICC Cricket World Cup, whereas lower-ranked teams must play a qualifying event—the ICC Trophy.

Rating individuals in team sports is a more complex task, primarily because of the team structure. Players

may play multiple positions. For example, baseball has specialist batters, pitchers, and fielders, and rating systems have been developed for each specialty, as Albert and Bennett (2003), Dewan (2006), and Keri (2006) discuss. In principle, comparing players with the same specialty is relatively straightforward, although several metrics have been suggested for this purpose. Albert (2009) provides a recent review of these metrics. However, complications arise when one wants to compare the contribution (to a team) of a player with a specialised role with another player who has a different specialised role. For example, comparing a specific batter's performance to that of a specific pitcher is not a simple task. Goldman and Kahr (2010) and James and Henzler (2002) discuss metrics that attempt such an analysis in baseball—to value players on a team regardless of their playing position. Cricket is similar in principle; researchers have explored methodologies for a common ratings system (Lewis 2005), as they have with baseball.

Soccer is more complex. Players have particularly varied roles, including goalkeepers, defenders, midfielders, and attackers (or strikers); each role requires a different sets of skills. However, each group also has subgroups. For example, midfielders can be disaggregated into wingers, defensive midfielders, and attacking midfielders. A team might not choose to play players at each position or role, depending on the strategy and composition of the opposing team. In addition to the complexity in team makeup, the interactions between the individuals being rated and their team members, and the interactions with the opposing players in the course of a game, make performance extremely dependent on external factors that the individual cannot control. For example, defenders and the goalkeeper (who is the last line of defence for stopping the opposition from getting the ball into the goal) who play on winning teams may have quiet games because their midfield and attack have been potent. In addition to these complex interactions, players' roles might change during a game; for example, fullbacks can play in both attacking and defensive roles, and would expect to vary their role depending on the state of the match, the tactics of the opposition, and instructions from the coach. In addition, some players might play a variety of positions during a game; therefore, rating a player who started

a game as a midfielder but finished it as a defender is difficult.

Despite these difficulties, the potential usefulness of a single rating system is clear. Teams can improve performance by identifying and adding undervalued players to their rosters, and the popularity of the sport can be enhanced by increasing debate among fans and pundits. In recent years, a wealth of data has become available, making it more attainable to rate all players using one scale. A rating system or competitor index could also be considered as a commercial product, and the information therein can be sold to the media. In this paper, we describe the development of such a system, the EA Sports Player Performance Index (PPI), the official player rating system of the PL (the top tier of soccer in England) and the Championship (the second tier of soccer in England). We describe the structure and construction of the system. We also highlight issues we encountered during the various stages of the consultation process and development, from initial contact to launch, and the challenges of balancing academic research with industry and corporate requirements. We then consider the values for players at the time of this writing, and conclude with a discussion that highlights opportunities for further development.

Description of the Development Team for the Index

The EA Sports PPI is a rating system for soccer players in England. The index is a partnership between the PL, the Football League (FL), Football DataCo (FDC), and the Press Association (PA). PL manages the top tier of soccer in England. FL manages the second tier of professional soccer in England and for certain Welsh teams playing in the English leagues. FDC is a joint venture company owned by the two professional leagues, which own the rights to data generated by the professional soccer competitions in England. PA is the official news agency for the United Kingdom. EA Sports PPI is the brand name for the index, which was known as the Actim index until 2010. The index forms part of a wider package of sports news information that is sold commercially around the world by PA acting as an agent for FDC. The development of the index began in 2003. Following an initial tender from interested academics, the

academic authors of this paper were fortunate to be given the opportunity to present ideas on how to construct an index for the clients (i.e., PL, FL, FDC, and PA), and were commissioned as consultants in 2004 to develop the player rating system.

Prior to formal meetings, the clients made it clear that the new index had to be based on sound mathematical reasoning; the media would use it to raise the Actim brand's profile. The clients and consultants agreed that the new index should: (1) be built on a statistical basis so that subjective opinions are not used when ranking players or judging the value of player activity (i.e., judging which players are more valuable); (2) be able to compare players from different positions; (3) be transparent so that nonstatisticians could understand how points are awarded; and (4) use a scale with which the general public would be comfortable—whole numbers for index scores were considered desirable. Furthermore, the clients required that the goals scored should be a direct component of the final index; they viewed the number of goals scored as too important a metric to be omitted from the index, although the conversion of scoring attempts into goals might be considered as occurring by chance (Pollard et al. 1977). After an initial one-year run, the consultants agreed to update the index to meet additional requirements to include points for players whose team did not concede a goal during a match (i.e., a clean sheet) and points for the player who makes the final pass to a player who scores a goal (i.e., an assist).

We should explain the rules of the game for the North American audience. They are essentially very simple: playing on a pitch of approximately 100 yards in length, two teams of 11 individuals each attempt to use their feet and heads (but not hands) to put the ball into the opponent's goal; after normal time, the team with the most goals wins; tournament rules can vary; some allow extra time to be played to determine winners and others allow ties (draws).

We also note that the relationship between academic consultants and industry clients is interesting. Various authors have examined the academic-industry consultation process, recently and in the past, and in both operations research (Dudley 1962, Ormerod 2002) and statistics (Marquardt 1979, Kenett and Thyregod 2006). During the consultancy project that we describe

in this paper, the most challenging aspects of the process for the academics were managing expectations, determining desirable properties of the developing index, and communicating the statistical nuances of the index (e.g., short-term variations). Where the consultant might like to be given *carte blanche* for the development of the solution, the client(s) might partly insist *a priori* on the form of the solution. The use of metrics for the number of goals scored, assists, and clean sheets was an example. The converse of this point is the acknowledgment that the client is perhaps better placed to consider the reactions of users of the solution and the nature of the solution. A compromise must ultimately be sought between complexity and simplicity, between the pure and the practical, and between the radical and the conservative. We elaborate on these experiences later in the text.

The Index

During the period of development, discussion, and evolution, all parties agreed that the final index would be a combination of individual component indices, each with its own properties, construction method, and emphasis. The individual subindices capture: (1) match contributions, (2) winning performance, (3) match appearances, (4) goals scored, (5) assists, and (6) clean sheets. As mathematical modellers, index 1 (i.e., match contributions) was of particular interest to us. However, as a result of high variability from match to match and from week to week, the clients felt that this needed stabilising using more standard (and simple) measures of performance.

We built the component indices using two seasons of PL data (2002–2003 and 2003–2004). The PA collected these data, and continue to do so, using a system called Football Live. This system provides real-time information on player actions and match events, essentially to provide ball-by-ball or kick-by-kick match reports on the Internet and on mobile platforms. A match reporter (formerly a professional player) at the match and a data logger at the PA base record the Football Live data; the reporter and data logger are in constant communication throughout a match via mobile phone link. Some match action or events (most notably passes and tackles) occur too quickly and at too great a frequency to

be captured effectively by the Football Live system. Postmatch video analysis captures these events. This data-capture process describes match events (i.e., contributions) and associates a named player with each event or contribution. Thus, the data contain information on match histories of player contributions. The first component index uses these data to the largest extent. We call this component index subindex 1 and describe its concept in the following subsection, leaving the mathematical exposition of its construction for the appendix. Descriptions of the other subindices then follow. The final ratings system consists of a weighted sum of six subindices.

Subindex 1: Modelling Match Outcome

McHale and Scarf (2007) describe a model that relates shots to player actions. The first subindex of the EA Sports PPI is based on a simpler version of this model. The flowchart in Figure 1 represents the levels that make up this index (considered from the perspective of the home team). We describe the model in the following paragraphs and the mathematical development of the subindex in the appendix.

We can think of the model in Figure 1, from right to left of the flow diagram, as follows. The result in any one game is a function of the goals for and goals against each team in the game. This function is deterministic; that is, if goals for and goals against are known, then the game result and hence the points awarded to each team (three points for a win, one for a draw (tie), and zero for a loss) are completely determined.

The goals for and goals against depend on the number of shots and the shot effectiveness of each team in the game. We interpret the second stage as a statement of probability, in line with Pollard et al. (1977), who suggest that goals can be interpreted as a random consequence of shots. In our model, the number of goals by each team is simply the number of shots each team has, multiplied by the probability that the shot results in a goal. We call this latter probability the shot effectiveness; it is the product of the attacking team's shot accuracy, the defending team's blocking capability, and the defending team's goalkeeping ability. The attacking team's shot accuracy is the probability that the home-team shot is on target; the defending team's blocking capability is the probability that the

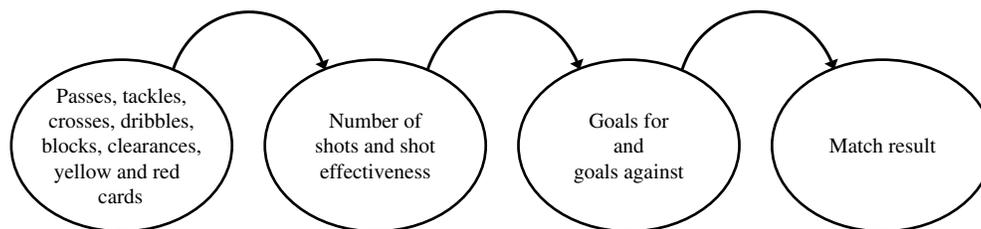


Figure 1: The flowchart depicts an influence diagram for our match outcome model (subindex 1): match result is a function of goals for and goals against that are determined by the number of shots and shot effectiveness of the two teams, which are determined by the actions of each team (e.g., passes and tackles).

away team blocks the home-team’s shot conditional on it being on target; the defending team’s goalkeeping ability is the probability that the away team saves the shot conditional on it being on target and not being blocked. This aspect of the model captures the effect of player contributions relating to shots (on and off target), blocks, and saves.

We model the number of shots for each team as independent Poisson random variables with means that we call the shot rates. The shot rate for a team is the expected number of shots for that team. We model this shot rate as a linear function of the number of player contributions not yet captured in the shot model (e.g., passes, crosses, dribbles, interceptions, and red and yellow cards). The linear form allows us to interpret the coefficient of a player contribution as the expected number of extra shots when the number of that contribution increases by one. We treat tackles differently in that we consider the ratio of tackles won to tackles lost as the relevant covariate; treating tackles won in a linear manner resulted in a negative coefficient; hence, a player would lose points for a successful tackle.

Given that the number of shots for the home team and against the home team are Poisson distributed and the shot effectiveness is the probability that a shot results in a goal, and assuming (as we do) that the shot effectiveness does not vary with the number of shots, the number of goals for and goals against the home team are also Poisson-distributed random variables. The shot process is effectively thinned to give the goal process. Win and draw probabilities for each team are then determined using the Poisson probabilities for goals for and goals against (e.g., $\text{prob}(\text{home win}) = \text{prob}(1, 0) + \text{prob}(2, 0) + \text{prob}(2, 1) + \dots$ in an

obvious notation). League points for a team resulting from the match outcome then follow; we award to a player for a single contribution X , the change in the expected number of points for his team when the number of contributions X changes by one: that is, the derivative of the expected number of points with respect to the number of contributions X . The derivative for each contribution can be determined from the double Poisson model for goals for and goals against. The derivative in each case is evaluated with other contribution types held at specific values. For contributions in the shot-rate part of the model, these are team-within-match specific values. This implies that for these contributions, the points awarded for a single contribution in the match do not vary within a team, but do vary between matches and between teams within a match. For contributions in the shot-effectiveness part of the model, specific values could be defined in a number of ways, as we discuss below.

We estimate the exact nature of the shot model from the data on past matches (i.e., 760 matches in 2002–2003 and 2003–2004 seasons). As one would expect, a relationship exists between the number of shots by a team and the player contributions. Indeed, shots increase as passes increase. Similarly, as red cards for the opposition increase, shots are expected to increase. These assertions are not altogether trivial. For example, is it the case that by passing more, you will create more shots on goal? Or is the relationship borne because the better teams, who have more shots, tend to pass the ball more? We do not attempt to establish cause and effect, but instead quantify the association between passes and shots.

As McHale and Scarf (2007) discuss, the relationship between the expected number of shots and

player actions is considered using a regression model. However, for ease of interpretation of the coefficients in the model, we use a simplification. We estimate the model for the expected number of shots by a team using ordinary least squares (not distinguishing between the home team and the away team). One could (and probably should) estimate the regression model allowing exactly for the Poisson nature of the dependent variable, and using a log-linear link function so that the logarithm of the expected number of shots is linear in the player contributions. Additional complexity can be added to the modelling by allowing for the dependence between the goals scores by the two teams in a match. However, our experiments revealed that the final index was no more intuitive with more complicated models for shots. A further advantage of not distinguishing between the home and away teams in this context is that it ensures that players are awarded the same number of points for actions, regardless of whether they were playing for the home or away team—a property deemed desirable by the clients. Table 1 shows the fitted model for expected number of shots.

The positive (and statistically significant) coefficient on passes, for example, confirms that teams that pass more tend to take more shots. Crosses have a large coefficient; this is not surprising because a cross into the opposition penalty area often leads to a goal-scoring opportunity. The largest coefficient here is for red cards; hence, our model implies that the most

important contribution to a team's shot rate is a red card (i.e., sending off) for an opposition player. Again, this is not surprising because teams that are down by one player tend to be dominated. However, red cards are not a major overall contributor to the index because they are relatively rare in comparison to other contributions. The signs of the coefficients are all as we would have expected. Where the evidence for a real covariate effect is weaker (i.e., larger p value) for red and yellow cards and clearances, a possible explanation is that these events tend to occur less frequently than others. A possible explanation of the weakness of the tackle ratio effect is that the current model does not ideally capture the influence of tackles. A large number of tackles won, for example, can indicate both a team being dominant (winning the ball back) and being dominated (desperate defending). Therefore, the effect of tackles is not straightforward.

Estimation of the probabilities in the shot-effectiveness part of the model can proceed in a number of ways. The first is to use game-by-game estimates. A consequence of this is that the shot effectiveness will vary from team to team and game to game; a team that scores no goals in a game will get no points for contributions in that game. The second is to use season-long home and away estimates of shot effectiveness. In this second approach, a home-team contribution will receive the same points regardless of the team and game. Other approaches would be to estimate shot-effectiveness probabilities by team and home-away status, but on a season-long basis, so that teams who consistently score more goals would get more points for contributions such as passes. We could ignore the home-away effect. For the final subindex, the clients deemed that it would be most desirable to hold shot-effectiveness probabilities constant. This decision was taken partly based on the idealistic principle that the points awarded to a player should not depend on the shot effectiveness of other players in the game, and partly in response to the higher variability in points awarded to a player for his actions in a single game using the other procedures for estimating the shot effectiveness. We also considered that players lose points for shots off target because this may seem harsh on strikers who are sometimes given the responsibility of shooting from

| Variables | Coef. | Std. err. | t -stat | p -value |
|-----------------------------|--------|-----------|-----------|------------|
| Crosses | 0.519 | 0.069 | 7.490 | 0.000 |
| Dribbles | 0.118 | 0.026 | 4.600 | 0.000 |
| Passes | 0.034 | 0.004 | 7.830 | 0.000 |
| Opposition interceptions | -0.024 | 0.009 | -2.520 | 0.012 |
| Opposition yellows | 0.253 | 0.134 | 1.890 | 0.059 |
| Opposition reds | 1.023 | 0.506 | 2.020 | 0.043 |
| Opposition tackle win ratio | -0.170 | 0.090 | -1.890 | 0.059 |
| Opposition cleared | -0.017 | 0.009 | -1.920 | 0.055 |
| Constant | 6.463 | 0.815 | 7.930 | 0.000 |

Table 1: The table depicts the results for a model of the shot rate as used in our index for rating player performance in a soccer match. For each contribution type, the coefficient, its standard error, and t statistic and p -value for test of significance of hypothesis, H_0 : coefficient equals zero are shown. Increasing the number of crosses and increasing the opposition's number of red cards are two ways in which a team can increase its expected number of shots.

difficult positions, inducing shots off target. However, because goals scored contributes to the overall index (see below), we retained the shots-on, shots-off balance in subindex 1 in the final implementation.

We note that we rescaled the points for saves to account for two features. First, a large proportion of saves can be regarded as routine saves. Second, only one player on the pitch can make a save. This is in contrast to any other action for which points are awarded (e.g., tackles or passes for which any player on the pitch can gain points). The Football Live data contained information on whether each save was identified as easy, normal, or difficult. We scale the points for saves by the proportion of saves that were identified as difficult. Alternatively, we might have awarded negative points to goalkeepers for goals conceded.

Although the model description above is from the point of view of the home side, away-team points for players may be calculated by determining the effect of away-team contributions on the expected number of points for the away team in a similar manner.

Finally, we rescale the points awarded for each contribution so that the total points over all players and games in a season is equal to the total league points in the final table (summed over teams). The points awarded to a player in a match I_1 is the sum over all contribution types of the product of the number of contributions by the player multiplied by the value of each contribution; this value is the expected change in league points for the team, given one additional contribution.

Other Subindices

Subindex 2: Points-Sharing Index

For each game, the construction of this subindex uses the number of minutes played by the player, the total number of minutes played by all players on his team, and the number of points the team won in that game. The points awarded to a player in a match on subindex 2, I_2 , is given by the product of the points won by the team and the ratio of the number of minutes played by the player to the total number of minutes played by the entire team, including substitutes. This subindex effectively shares out the points won

by the team among the players who played according to how long they spent on the pitch. Note that the total points awarded over the season (summed over the players) will equal the total league points in the final league table (both overall and by the team). Winning players will lead this subindex; it will have a mix of strikers, midfielders, defenders, and goalkeepers because only time on the pitch and results are considered. Points allocated to players over a single weekend are highly variable; one reason is that many players on losing sides will receive zero points, perhaps in spite of good individual performances.

Subindex 3: Appearance Index

This subindex divides the number of points won by all teams in the league among the players according to how many minutes they played. This subindex is similar to subindex 2; however, this index does not consider results—points are awarded for playing. Subindex 3 rewards players who play a high proportion of games during the season (i.e., team stalwarts). The average number of points won in any one game by any one team is 1.34 (calculated on the basis of data from the past 10 seasons); thus, the points awarded to a player in a match on subindex 3, I_3 , is given by the ratio of the number of minutes played by the player to the total number of minutes played by the entire team, including substitutes, multiplied by 1.34. The factor 1.34 is included here so that the points for this index are on the same scale as the points for subindices 1 and 2. For subindex 3, it follows that the total points over all players and games in a season is equal to the expected total league points in the final table (summed over teams).

The concept of rewarding players for appearances is not trivial. For example, teams play more high-profile games than ever before; as such, the physical demands placed on players are extremely high. A common tactic used by coaches is to rest star players for certain games—either less important games or perhaps games against lesser opposition. This squad-rotation policy is a complication for the index because the best players are not always chosen when they are available; therefore, the index may not identify the true quality of the top players because their performance is not measured against the full spectrum of opposition ability.

Subindex 4: Goal-Scoring Index

This subindex awards points to players based on their goals scored. The specific number of points awarded has been calculated by converting goals into points. At the time of the analysis, over the previous 12 seasons, 12,492 goals were scored in the top division and 12,985 league points were won. Therefore, we estimate how many points one goal is worth as $12,985/12,482 = 1.039$ league points. Thus, on this index, a player receives 1.039 points for each goal the player scores. The points awarded to a player in a match for goals, I_4 , is then simply the points per goal multiplied by the number of goals.

Subindex 5: Assists Index

Analogous to the goals-scored index, each player who provides an assist gets 1.039 points for that assist. This puts subindex 5 on the same scale as the other subindices. The points awarded to a player in a match for assists, I_5 , is then the points per assist multiplied by the number of assists.

Subindex 6: Clean-Sheets Index

To maintain the balance of the overall index, we take the total points awarded for clean sheets to be equal to the total points for assists. In the seasons analysed (i.e., 760 games in 2002–2003 and 2003–2004), there were 1,069 assists and 399 clean sheets. Therefore, we calculate the points per clean sheet as $(1.039 \times 1,069)/399 = 2.784$. We now need to divide these points among the entire team that achieved the clean sheet; thus, the points for the clean sheet are not awarded to a single player. Note that the total points for a clean sheet is close to, but not quite, three; this reflects that a clean sheet implies a win—but for a 0–0 result.

The data collected by PA Sport include several defensive actions: blocks, clearances, tackles won, interceptions, and saves. A weighted sum of these actions by position provides a distribution of defensive actions across a team (see Table 2). Thus, following this distribution for the points for a clean sheet, of the 2.784 points available, a goalkeeper should get $0.21 \times 2.784 = 0.585$; each defender, midfielder, and striker should receive a similar number of points. Thus, if a team achieves a clean sheet, the subindex points awarded to a player, I_6 , are 0.585

| Position | Percent of total defensive actions (%) |
|-----------------------------|--|
| Goalkeeper | 21 |
| Defender (per defender) | 13 |
| Midfielder (per midfielder) | 5 |
| Striker (per striker) | 3 |

Table 2: In a soccer match, some player actions can be described as defensive (i.e., blocks, clearances, tackles won, interceptions, and saves). The table shows the percentage of total defensive actions made per position (based on four defenders, four midfielders, and two strikers per team); we calculated the values from the 2002–2003 and 2003–2004 seasons and used them to distribute the points for a clean sheet among the team.

(goalkeeper), 0.364 (defender), 0.150 (midfielder), and 0.071 (striker). Player positions are defined as at season start (although player position defined for each match would facilitate a better assignment of clean-sheet points).

Final Index

The final index is a weighted sum of the points achieved on each subindex:

$$I = 100 \times (0.25I_1 + 0.375I_2 + 0.125I_3 + 0.125I_4 + 0.0625I_5 + 0.0625I_6).$$

A pure index would comprise only subindex 1 or a variation of it. However, this index does not provide a practical solution to the problem. Essentially, it has too much noise to allow its use for reporting on a weekly basis. Also, subindex 1 is perhaps too radical given the perception among pundits and fans that winning (and hence goals, assists, and clean sheets) determines the best players. Part of this perception has become firmly established because of the use of such metrics in fantasy soccer games. The Telegraph fantasy football website (<http://www.fantasyfootball.telegraph.co.uk>) provides details about fantasy soccer. Also, the role of chance in the game is perhaps too strong in subindex 1. Although uncertainty of outcome is an important factor in maintaining interest in the game, observers of the game believe that systematic effects are more dominant than they are. For example, a manager who loses x games in a row is blamed and replaced (Audas et al. 1999, De Schryver and Eisinga 2009). Thus, the final index comprises

a weighted sum of the six subindices. We experimented with weighting combinations to find the combination that most satisfied the client’s desire to have an index that produced stable ratings, but that was varied enough from week to week to generate discussion among fans, the media, and pundits. In this way, subindices 2–6 can be thought of as stabilizing components that regulate the variability inherent in subindex 1. Thus, the choice of weighting for the subindices reflects the compromise between purity and practicality, and between revolution and conformity.

Each subindex possessed the desirable property that total points available were approximately equal to league points available in the season; therefore, a player’s points might be interpreted as his share of the points won by the team. Despite this property, we rescaled the final index; when the clients considered the likely audience for the final index, they decided that it was undesirable to report rating points to two decimal places. The index is intended to be a talking point among fans, which can drive theoretical projections such as the components of a super team of top index players in each playing position. Generally, debate is a large factor in the life of any soccer fan; such debate takes place in homes, at work, in public houses, and in cafes, and also on the Internet as fans embrace the world of blogging and tweeting to settle the argument of who is the best player in their league or on their team.

The Index for the 2008–2009 Season

In this section, we present subindex 1 and the final EA Sports PPI (then known as the Actim index) for the 2008–2009 season of the PL competition. Table 3 shows the subindex 1 final totals. Several observations can be made from the list in Table 3. First, of the top 20 players, 7 are goalkeepers and 11 are defenders. This is a consequence of the deduction of points for off-target shots. The inclusion of goals scored and assists in the final index (see Table 4) moderates this outcome. The final standings in Table 4 have a good mix of players from different positions: six strikers, five midfielders, four defenders, and five goalkeepers. One of the original desirable qualities of an index was to include players who are not necessarily from the

| Name | Team | Position | Subindex 1 points |
|--------------------|----------------------|------------|-------------------|
| Mark Schwarzer | Fulham | Goalkeeper | 7.29 |
| Gareth Barry | Aston Villa | Midfielder | 7.06 |
| Sol Campbell | Portsmouth | Defender | 6.86 |
| Gary Cahill | Bolton Wanderers | Defender | 6.70 |
| J. Lloyd Samuel | Bolton Wanderers | Defender | 6.63 |
| Robert Green | West Ham United | Goalkeeper | 6.45 |
| Heurelho Gomes | Tottenham Hotspur | Goalkeeper | 6.43 |
| David James | Portsmouth | Goalkeeper | 6.40 |
| Scott Carson | West Bromwich Albion | Goalkeeper | 6.39 |
| Nicolas Anelka | Chelsea | Striker | 6.33 |
| Richard Dunne | Manchester City | Defender | 6.30 |
| Andrew O’Brien | Bolton Wanderers | Defender | 6.26 |
| Jussi Jaaskelainen | Bolton Wanderers | Goalkeeper | 6.20 |
| Sylvain Distin | Portsmouth | Defender | 6.12 |
| Thomas Sorensen | Stoke City | Goalkeeper | 6.10 |
| Matthew Upson | West Ham United | Defender | 6.03 |
| Ryan Shawcross | Stoke City | Defender | 5.85 |
| Danny Collins | Sunderland | Defender | 5.81 |
| Michael Turner | Hull City | Defender | 5.54 |
| Anton Ferdinand | Sunderland | Defender | 5.47 |

Table 3: For our player rating index based on a model of match outcome (subindex 1), the table shows the top 20 players for the 2008–2009 season. The list is dominated by goalkeepers and defenders. The shortage of strikers is because of the negative impact on subindex 1 of off-target shots.

top teams. One might consider this as having been successfully achieved given the appearance of Mark Schwarzer of Fulham, Joleon Lescott and Tim Howard

| Name | Team | Position | Index score |
|--------------------|-------------------|------------|-------------|
| Nicolas Anelka | Chelsea | Striker | 751 |
| Dirk Kuyt | Liverpool | Midfielder | 655 |
| Frank Lampard | Chelsea | Midfielder | 653 |
| Dimitar Berbatov | Manchester United | Striker | 586 |
| Gabriel Agbonlahor | Aston Villa | Striker | 575 |
| Jose Reina | Liverpool | Goalkeeper | 573 |
| Gareth Barry | Aston Villa | Midfielder | 564 |
| Jamie Carragher | Liverpool | Defender | 552 |
| Steven Gerrard | Liverpool | Midfielder | 540 |
| Nemanja Vidic | Manchester United | Defender | 540 |
| Cristiano Ronaldo | Manchester United | Striker | 517 |
| Petr Cech | Chelsea | Goalkeeper | 513 |
| John Terry | Chelsea | Defender | 501 |
| Robbie Keane | Tottenham Hotspur | Striker | 482 |
| Edwin Van der Sar | Manchester United | Goalkeeper | 481 |
| Neves Denilson | Arsenal | Midfielder | 477 |
| Mark Schwarzer | Fulham | Goalkeeper | 477 |
| Joleon Lescott | Everton | Defender | 476 |
| Tim Howard | Everton | Goalkeeper | 472 |
| Kevin Davies | Bolton Wanderers | Striker | 468 |

Table 4: The table shows the end-of-season final index totals for the 2008–2009 season for the top 20 players. The composition is a good mix of players from all positions and all teams.

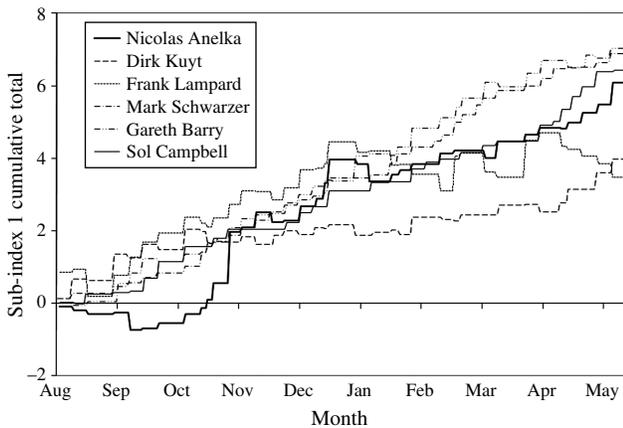


Figure 2: The graph shows cumulative totals over time for the top three players for the 2008–2009 season (Schwarzer, Barry, and Campbell) according to our subindex 1 and the top three players from the final index (Anelka, Kuyt, and Lampard). The negative impact of shots off target is particularly evident for Nicolas Anelka in the early months of the season. One can see how poor performances result in negative index totals.

of Everton, and Kevin Davies of Bolton Wanderers. Although players in the top teams (i.e., Manchester United, Liverpool, and Chelsea) lead the index, player rotation, in which the top teams with large squads rest their players, means that they do not dominate it.

Figure 2 shows the variability in subindex 1, the pure contributions index. Although Nicolas Anelka finished at the top of the overall index, he had a negative total on subindex 1 for the first two months of the season. The goalkeeper (Schwarzer)

and defensive midfielder (Barry) are much more stable on this subindex. Lampard, infamous for shooting off target, is highly variable. The inclusion of the other subindices in the final index moderates this variability.

Following the 2010 World Cup in South Africa, comparing a hypothetical England team picked using the index with the actual England team chosen is interesting (see Table 5). The team in Table 5 is not identical to the team England chose to field in the opening game of the World Cup. One omission is notable: Steven Gerrard of Liverpool is missing. Gerrard was injured for a large part of the season; therefore, he did not receive as many points as he would have if he had been fit. This raises an important point regarding the index—players can only obtain rating points if they play; therefore, the index provides a rating of the players who are playing the best soccer and not necessarily a rating of the best players. We leave it to the fans to produce a per-minute-played index.

Discussion

The EA Sports PPI is a performance rating system for players in the top divisions of soccer in England. It is now in its seventh season of use. The index was developed by the authors in collaboration with the soccer leagues and a leading news media company, and is a commercial news product that is sold to news media outlets worldwide. It is published weekly

| Actim Index ranking | Name | Team | Position | Index score | Actual player |
|---------------------|-----------------|-------------------|------------|-------------|----------------|
| 3 | Wayne Rooney | Manchester United | Striker | 482 | Wayne Rooney |
| 12 | Jermain Defoe | Tottenham Hotspur | Striker | 422 | Emile Heskey |
| 9 | Frank Lampard | Chelsea | Midfielder | 433 | Frank Lamard |
| 16 | James Milner | Aston Villa | Midfielder | 376 | James Milner |
| 20 | Ashley Young | Aston Villa | Midfielder | 370 | Aaron Lennon |
| 21 | Gareth Barry | Manchester City | Midfielder | 368 | Steven Gerrard |
| 19 | John Terry | Chelsea | Defender | 371 | John Terry |
| 24 | Leighton Baines | Everton | Defender | 357 | Glen Johnson |
| 27 | Ashley Cole | Arsenal | Defender | 348 | Ashley Cole |
| 40 | Ryan Shawcross | Stoke City | Defender | 325 | Ledley King |
| 44 | Joe Hart | Birmingham City | Goalkeeper | 322 | Robert Green |

Table 5: The table shows a hypothetical England team chosen from the highest-ranked players in their playing position as of March 12, 2010 (column 2), and the actual team that played the first game in the 2010 World Cup finals (final column). The two lineups are closely matched, especially when we consider that Steven Gerrard and Rio Ferdinand were injured for a large part of the season; thus, they were unable to accumulate points.

by the PL. The index is also used as part of the PL Fantasy Football game, and is published in match-day programmes and newspapers around the globe. The index was the first to rate individual players in a team sport using a single common score, regardless of a player's position. This single-rating score was a particular challenge, which we addressed during the index development. We were able to meet this objective partly because soccer players multitask; although players have specialities, roles can change subtly within a game, and all players have the opportunity to make a range of winning contributions, for example, passes, tackles, and shots. Goalkeepers also make contributions other than saves, although infrequently.

The PL and Championship in England use the index in their annual competitions, and the PL uses it in its Fantasy Football competition. The index is updated weekly during the season; the list of current top 100 players in the PL is published on the PL website (<http://www.premierleague.com/page/ActimIndex>) each Tuesday. Lists of best defenders, midfielders, attackers, and goalkeepers are also published, as is the EA sports team of the week. The team of the week comprises the players who score mostly highly on the index during the previous week, subject to the constraint that the team must include one goalkeeper, four defenders, four midfielders, and two attackers. The team of the week consequently changes on a week-by-week basis. The overall index evolves more slowly, given that a player's index points accumulate over the season; an injury or layoff can lead to a significant reversal of fortune for a player. The ratings and rankings are also published in newspapers and match programmes, and are debated by observers and commentators worldwide. An Actim index blog (<http://www.actimindex.wordpress.com/the-actim-index>) is also published (note that accessing the blog requires user registration).

The availability of detailed data collected both during and after matches regarding players' contributions and the notion that such data capture the essence of the *beautiful game* essentially drove the development of the index. However, we found that an index based purely on player contributions was too volatile for weekly publication. The clients for the index have a more deterministic focus, and the final

index comprises a weighted average of those contributions that directly influence success (e.g., league points won, goals) and those contributions that statistically influence success (e.g., passes, tackles). In this latter case, we mean those contributions that appear, on the basis of our statistical analysis, to be associated with goals scored.

As data relating to player contributions become more sophisticated and more representative of the game, we expect further development of the index. For example, the collection of pass origin and finish is now feasible, as is the collection of data relating to sequences of passes. Such data could allow complex interactions between player actions to be modelled in a way that is not currently considered in the index construction. Additionally, for season 2010–2011, the PL is introducing a player tracking system that will provide distance covered by all moving objects (including the ball) engaged on the field of play, coupled with event-time recording to provide additional calculations relating to events in the game (e.g., speed, distance, acceleration, deceleration, force of pass). The future of sports statistics is becoming ever more sophisticated.

Appendix

In the model development, we use the following notation for a game.

H_{PTS} : number of points for the home side (0, 1, or 3), depending on outcome.

A_{PTS} : number of points for the away team.

H_G : number of goals for the home team.

A_G : number of goals for the away team (goals against the home team).

H_X : number of contributions X by the home team— X may refer to passes, tackles, etc.

A_X : number of contributions X by the away team.

H_S, A_S : number of shots (home, away)

$E\{H_S\}$: home-team expected number of shots per game (i.e., shot rate), respectively.

$p_{HG|HS}, p_{AG|AS}$: probability of a home (away) goal given a shot by the home (away) team (shot-effectiveness).

$p_{HSAVE|AS}$: probability of a home save given an away-team shot.

$p_{HBO|AS}$: probability of a home-team block given an away-team shot.

$p_{HS_ON|HS}$: probability of a home-team shot on target given a home-team shot.

In terms of the three stages discussed following Figure 1, the model structure is as follows. The points for the home

side, H_{PTS} , depend on the goals for and goals against, so that

$$H_{PTS} = \begin{cases} 3 & \text{if } H_G > A_G \\ 1 & \text{if } H_G = A_G \\ 0 & \text{if } H_G < A_G \end{cases},$$

because teams are awarded three points for a win, one point for a draw, and zero points for a loss. At the next level in the hierarchy, we model the shot effectiveness and the shot rate. For the shot effectiveness, we assume

$$p_{HG|HS} = (1 - p_{AB|HS})(1 - p_{ASAVE|HS})p_{HS_ON|HS}, \quad (1)$$

where $p_{AB|HS}$ is the probability that the away team blocks the home team's shot, $p_{ASAVE|HS}$ is the probability that the away team saves the shot, and $p_{HS_ON|HS}$ is the probability that the home-team shot is on target. For the away team we have

$$p_{AG|AS} = (1 - p_{HB|AS})(1 - p_{HSAVE|AS})p_{AS_ON|AS}. \quad (2)$$

The shot effectiveness is the probability of a goal given a shot; this is assumed to be the product of the probabilities that the shot is on target ($p_{AS_ON|AS}$), is not blocked ($1 - p_{HB|AS}$), and is not saved ($1 - p_{HSAVE|AS}$).

We model the goals for and goals against as independent Poisson-distributed random variables, so that $H_G \sim \text{Poisson}(p_{HG|HS}E\{H_S\})$, where $p_{HG|HS}$ is the effectiveness of the home-team (away-team) shots (shot effectiveness) and $E\{H_S\}$ is the expected number of home-team (away-team) shots per game. We model the expected number of shots (i.e., the shot rate) as a function of other aspects of play such as passes, dribbles, etc. Thus,

$$E\{H_S\} = f(H_X, H_Y, \dots, A_U, A_V, \dots), \quad (3)$$

and similarly for the away team. H_X is, for example, the number of contributions of type X (e.g., passes) by the home side and A_U is the number of contributions of type U (e.g., interceptions) by the away side. It is natural to use the exponential function for the function in Equation (3) above, although it is simpler to assume that

$$E\{H_S\} = \alpha_0 + \alpha_X H_X + \alpha_Y H_Y + \dots + \beta_U A_U + \beta_V A_V + \dots. \quad (4)$$

With this form, we can interpret the coefficient α_p , for example, as the expected number of extra shots when the number of passes increases by one.

Denoting the expected number of points for the home side by $E\{H_{PTS}\}$, the number of points we award to a player for a single contribution X is the change in the expected

number of points for his team when the number of contributions X changes by one, that is, $\partial E\{H_{PTS}\}/\partial H_X$. Noting that the double Poisson model implies

$$\begin{aligned} E\{H_{PTS}\} &= (3 \text{ points}) \times (\text{probability of a win}) + (1 \text{ point}) \\ &\quad \times (\text{probability of a draw}) \\ &= \exp(-\lambda_{HG} - \lambda_{AG}) \left\{ 3 \sum_{i=0}^{\infty} \sum_{j=i+1}^{\infty} \frac{\lambda_{HG}^i \lambda_{AG}^j}{i!j!} + \sum_{i=0}^{\infty} \frac{\lambda_{HG}^i \lambda_{AG}^i}{i!i!} \right\}, \end{aligned} \quad (5)$$

where $\lambda_{HG} = p_{HG|HS}E\{H_S\}$ and $\lambda_{AG} = p_{AG|AS}E\{A_S\}$, all the derivatives, $\partial E\{H_{PTS}\}/\partial H_X$, can be determined from the equation above and Equations (1), (2), (4), and (5). We omit these derivatives for brevity, but note that for each contribution, the derivative is evaluated with all other contribution types held at their match and team-within-match specific values. Although the model description above is from the point of view of the home side, away-team points for players may be calculated by determining the effect of away-team contributions on the expected number of points for the away team, $\partial E\{A_{PTS}\}/\partial A_X$, in a similar manner.

Estimation of $p_{HG|HS}$ and $p_{AG|AS}$ can proceed in a number of ways. The first is to use game-by-game estimates, so that $\hat{p}_{HG|HS} = H_G/H_S$ and $\hat{p}_{AG|AS} = A_G/A_S$. The second is to use season-long home and away estimates of shot effectiveness, that is, $\hat{p}_{HG|HS} = \bar{H}_G/\bar{H}_S$ and $\hat{p}_{AG|AS} = \bar{A}_G/\bar{A}_S$, where \bar{H}_G and \bar{H}_S are the season-long arithmetic means of the number of home goals and home shots per game, respectively (and similarly for away goals and shots). Another approach would be to estimate $p_{HG|HS}$ and $p_{AG|AS}$ by team and home-away, but on a season-long basis. The home-away effect can be ignored by setting

$$\hat{p}_{HG|HS} = \hat{p}_{AG|AS} = (\bar{H}_G + \bar{A}_G)/(\bar{H}_S + \bar{A}_S).$$

The points awarded to a home-team player i in a match, I_1 , is then $\sum_X H_{i,X} \{\partial E\{H_{PTS}\}/\partial X\}$, where $H_{i,X}$ is the number of contributions of type X by player i ; the formulation for away-team players is similar.

For subindex 2, the points awarded to player i in match t is given by

$$I_2 = \frac{\text{mins}_{it}}{\sum_i \text{mins}_{it}} \times \text{points}_t,$$

where mins_{it} is the number of minutes player i was on the pitch in game t , $\sum_i \text{mins}_{it}$ is the total number of minutes that all players on the team of player i (including player i) were on the pitch (approximately 90 minutes \times 11 players) in game t , and points_t is the league points won by the team of player i in match t , being equal to 3, 1, or 0 according to the outcome of game t .

For subindex 3, the points awarded to player i in match t are given by

$$I_3 = \frac{\text{mins}_{it}}{\sum_i \text{mins}_{it}} \times 1.34.$$

The other subindices are described in the main text.

Acknowledgments

The authors acknowledge the contribution of several people in developing the Actim index: Martin Axford (PL), Chris Buckley (ex PA), Neil Chambers (PA), Andy Elliot (ex PA), Guy Hollis (PA), Phil Lines (PL), Paul Molnar (PL), and Brian Phillipotts (ex PL). They also thank two anonymous referees and the associate editor for their suggestions, which have improved the paper.

References

- Albert, J. 2009. Exploring pitch F/X data. R. Koning, P. Scarf, eds. *Proc. 2nd Internat. Conf. Math. Sport*, the Institute of Mathematics and Its Applications, London, UK, 10–18.
- Albert, J., J. Bennett. 2003. *Curve Ball*. Springer, New York.
- Audas, R., S. Dobson, J. Goddard. 1999. Organizational performance and managerial turnover. *Managerial Decision Econom.* **20**(6) 305–318.
- De Schryver, T., R. N. Eisinga. 2009. The application of piecewise linear models to management turnover decisions in Dutch soccer clubs. R. Koning, P. Scarf, eds. *Proc. 2nd Internat. Conf. Math. Sport*, the Institute of Mathematics and Its Applications, London, 159–164.
- Dewan, J. 2006. *The Fielding Bible*. ACTA Sports, Chicago.
- Dudley, N. A. 1962. Operational research and the universities. *J. Oper. Res. Soc.* **13**(1) 81–86.
- Goldman, S., C. Kahrl. 2010. *Baseball Prospectus 2010: The Essential Guide to the 2010 Baseball Season*. John Wiley & Sons, Inc., Hoboken, NJ.
- Grinblatt, M., S. Titman. 1989. Portfolio performance evaluation: Old issues and new insights. *Rev. Financial Stud.* **2**(3) 393–421.
- James, B., J. Henzler. 2002. *Win Shares*. STATS Publishing Inc., Northbrook, IL.
- Kenett, R., P. Thyregod. 2006. Aspects of statistical consulting not taught by academia. *Statistica Neerlandica* **60**(3) 396–411.
- Keri, J. 2006. *Baseball Between the Numbers*. Basic Books, New York.
- Lewis, A. J. 2005. Towards fairer measures of player performance in one-day cricket. *J. Oper. Res. Soc.* **56**(7) 804–815.
- Lewis, M. M. 2003. *Moneyball: The Art of Winning an Unfair Game*. W. W. Norton and Co., New York.
- Marquardt, D. W. 1979. Statistical consulting in industry. *Amer. Statistician* **33**(3) 102–107.
- McHale, I. G., S. M. Davies. 2007. Statistical analysis of the FIFA world rankings. R. Koning, J. Albert, eds. *Statistical Thinking in Sport*. Chapman and Hall, Boca Raton, FL, 77–90.
- McHale, I. G., D. Forrest. 2005. The importance of recent scores in a forecasting model for professional golf tournaments. *IMA J. Management Math.* **16**(2) 131–140.
- McHale, I. G., A. Morton. 2011. A Bradley-Terry type model for forecasting tennis match results. *Internat. J. Forecasting* **27**(2) 619–630.
- McHale, I. G., P. A. Scarf. 2007. Modelling soccer matches using bivariate discrete distributions with general dependence structure. *Statistica Neerlandica* **61**(4) 432–445.
- Ormerod, R. J. 2002. On the nature of OR: Taking stock. *J. Oper. Res. Soc.* **53**(5) 475–491.
- Pollard, R., B. Benjamin, C. Reep. 1977. Sport and the negative binomial distribution. S. P. Ladany, R. E. Machol, eds. *Optimal Strategies in Sports*. North Holland, New York, 118–195.
- Swamidass, P. M., W. T. Newell. 1987. Manufacturing strategy, environmental uncertainty and performance: A path analytic model. *Management Sci.* **33**(4) 509–524.