VALUING GOALS, SAVES AND DUELS IN SOCCER

INTRODUCTION

Soccer constitutes the largest majority in global market share at 43%, followed by American football at 13% and baseball at 12%[1]. Thus, putting dollar amounts on certain actions in soccer is quite useful for investors and gamblers alike. In this report, we try to value certain actions in soccer. Some of the questions we try to answer are: How much is scoring a goal worth? How much is saving a goal worth? How much is winning a one-on-one duel worth? These questions may help teams better value certain actions and hence amplify those that are worth the most. To answer these questions, we analyze game data for the 2011-2012 season of the Premier League, as well as wage data. We then generalize our findings for a complete Premier League season to analyze the value of player performance in later seasons, as well as in other competitions.

DATA

Opta Sports, a London-based sports data company, produced the dataset used in the analysis. The dataset contained Opta's Classic Data feed from the 2011-2012 English Premier League season, and each row of the dataset consists of an individual's statistics in a single game. The dataset contains 184 columns of data, allowing for extensive analysis of all aspects of a soccer player's performance[2]. Opta Sports has specific event definitions that correspond to the statistics seen in our dataset. The following events were used in our analysis:

GOALS: Goals were used to asses the performance of strikers in our data set. Overall shot frequency, on target and off target was also used. It was assumed that the sole purpose of strikers is to score goals, although in reality, they could have other parallel roles such as playmaking and withdrawing defenders.

DUELS: A duel is a 50-50 contest between two players of opposing sides in the match. For every duel won, there is a corresponding duel lost depending on the outcome. We choose this metric for defenders over tackles. A tackle is awarded if a player wins the ball from another player who is in possession. We decided that duels are a better metric for defenders because some great defenders never steal the ball from the opponents, but excel by being in good positions and winning the important duels they must win to prevent attacks from resulting in goals.

SAVE: A goalkeeper preventing the ball from entering the goal with any part of his body. Goalkeepers have the smallest subset of statistics that pertain to them so saves were used as the only metric for goalkeepers in the analysis.

In addition to the game data, all monetary values used are sourced from transfermarkt.com. Throughout the analysis, market value of the player is used to determine a player's monetary worth, specifically the market value of the given player in June 2012, at the end of the 2012 English Premier League season.

[1] https://www.atkearney.com/documents/10192/6f46b880-f8d1-4909-9960-cc605bb1ff34

[2] https://datahub.io/dataset/uk-premier-league-match-by-match-2011-2012

METHODS

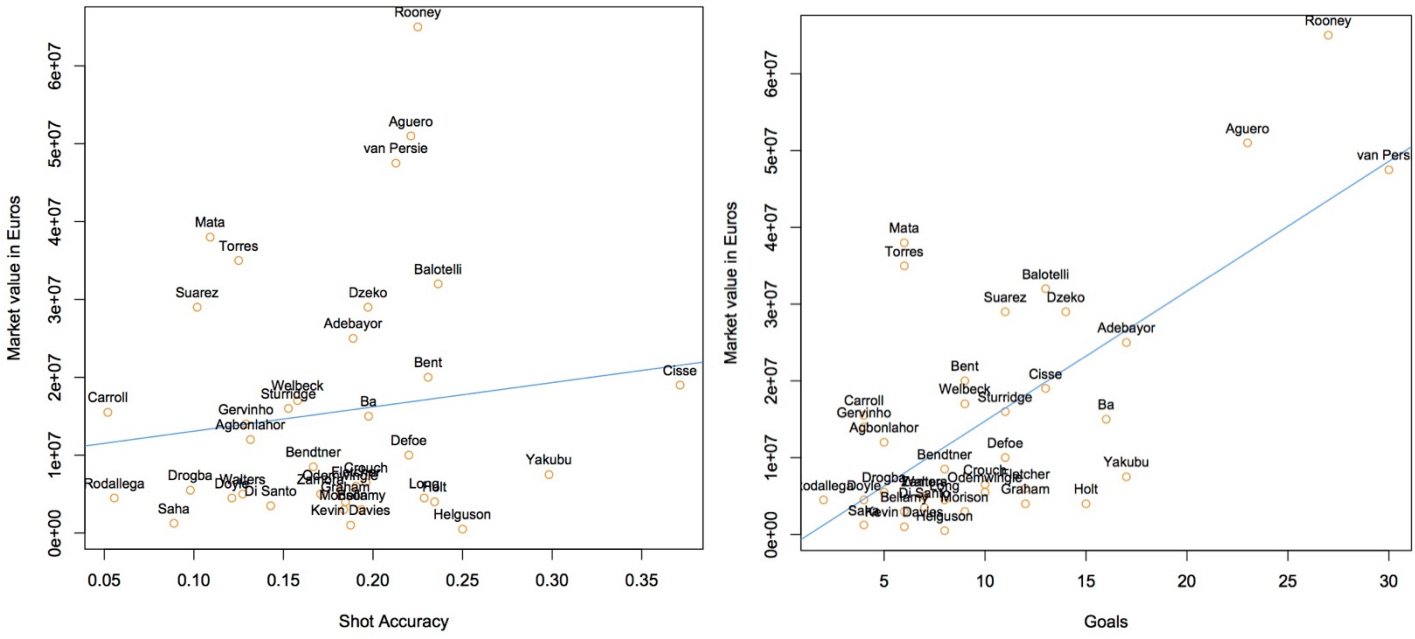
When evaluating players, regression to the mean (RTTM) was used to determine a player's true talent compared to the observed statistic. When evaluating players as possible transfer acquisitions in a competitive soccer league, we believe it is important to consider a true talent metric that accounts for the observed statistics of all soccer players' in the English Premier League. Sometimes pure observed percentage can be misleading, and to account for this, we subset the data to include only the most active players at each position and furthermore, regress towards the mean to determine a player's true talent compared to the rest of the players in the league at the same position. All basic analysis was done in R.

RESULTS

The result of our analysis is decomposed by position. Key actions of strikers, defenders and goalkeepers will be assessed and compared to each other in their values.

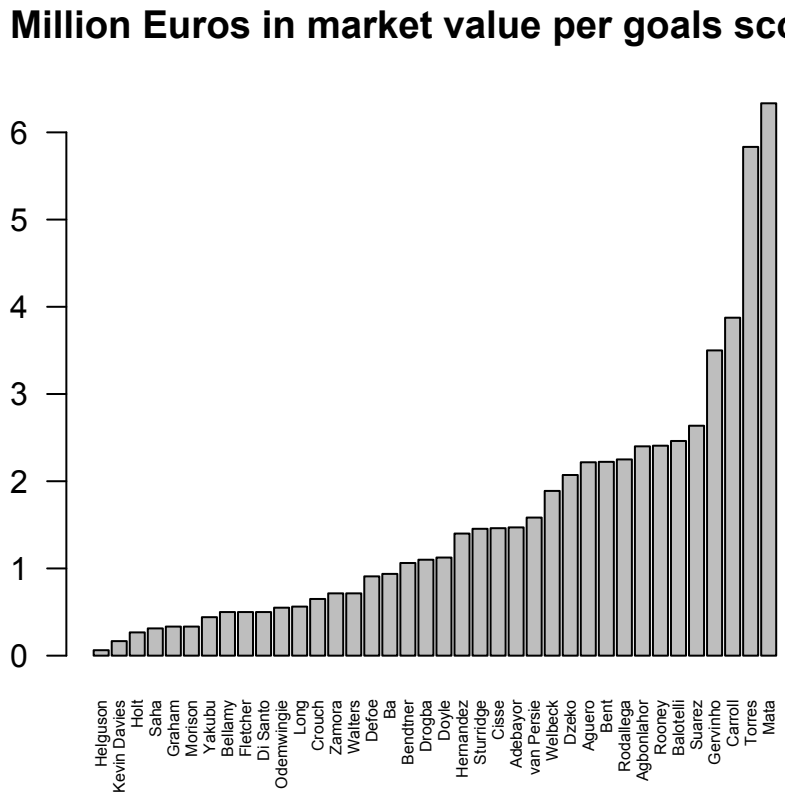
STRIKERS

From a top-down approach, the aim of any team in a soccer match is to win, and a team may not win unless they score at least one goal. Thus, it is evident to put a value on goals, as some striker contracts involve clauses for an extra bonus for each goal scored. Since most goals are scored by strikers, only players in such position were considered in this analysis with the objective to put a value on any goal scored. However, another thesis is to correlate player wages with their shot accuracy. When we plot the market value of the strikers with more than 30 shots in the season versus their goal count and shot accuracy, we get the following graphs:



From the graphs above, we calculated a correlation of 0.128 between market value and shot accuracy, or conversion rate, but obtained a correlation of 0.677 between market value and goals scored per season. One may argue that players that are highly paid are more likely to be in better teams that have better playmakers, and hence find it easier to score goals. However, this is a secondary factor in the correlation between wages

and goals scored, as the most direct factor is the inherent added value of scoring goals. Since goals scored was established as a better predictor of market value, we plotted the amount of money in market value per goal scored, and obtained the following profile:



The market value of a goal in the Premier League season of 2011-2012 was found to be 1.562m euros. This figure was close to other estimates in the literature. One study divided each club's annual income by the number of goals the club scored as a measure of the value of one goal. The average figure reported for the 2011-2012 Premier League season was 1.182m euros, which is in the range of our estimate here[3].

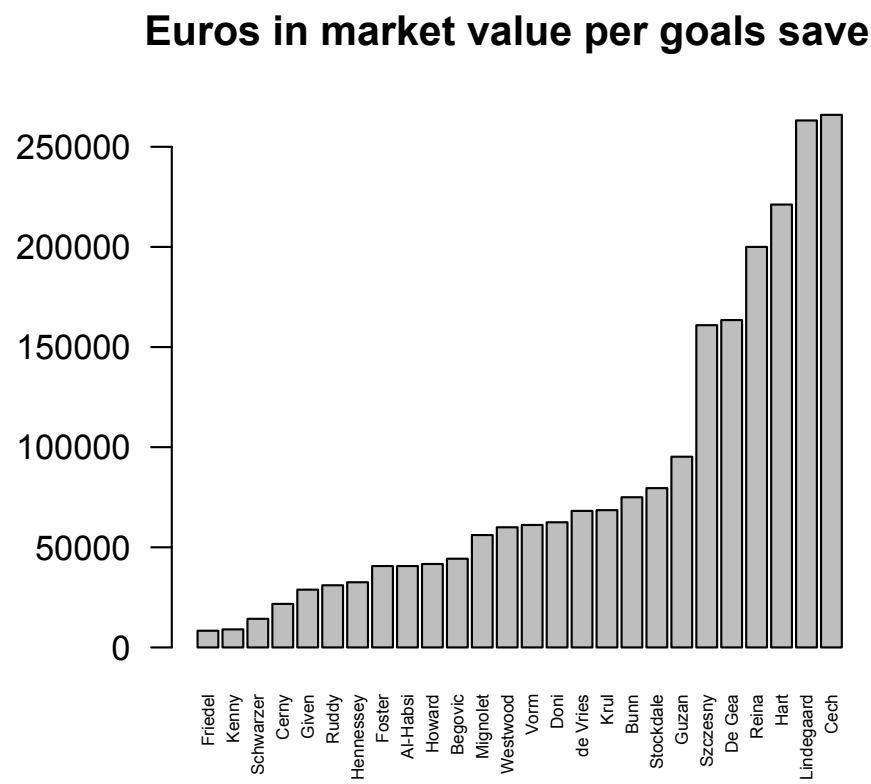
GOALKEEPERS

On the other side of the post lay goalkeepers. Unlike strikers, goalkeepers have more defined roles, and a single purpose: To not let goals go in. To assess their best performance indicator, we tested the correlation between the market value of goalkeepers that have played more than 180 minutes and a few metrics, which are outlined in the table below:

Performance indicator	Correlation to market value
Saves per season	0.301
Save percentage	0.183
Save Talent[4]	0.257

[3] http://www.sportingintelligence.com/2014/02/24/premier-league-goals-worth-915k-each-last-season-and-1-5m-now-240201/
[4] Calculated using the RTTM method outlined in the methods section

The best metric to correlate with goalkeeper wages was their saves per season, out of the ones tested, although the correlation is pretty low. Given this information, we also assess the value of a save using the same method as before. We then obtain the following distribution for the subset of goalkeepers considered:



From the distribution above, we notice an interesting distinction between goalkeepers of the "Big Five" in the premier league, and others. Removing goalkeepers of the "Big Five[5]", we obtain the following correlations:

Performance indicator	Correlation to market value	Correlation to market value without "Big Five" goalkeepers
Saves per season	0.301	0.678
Save percentage	0.183	0.183
Save Talent	0.257	0.209

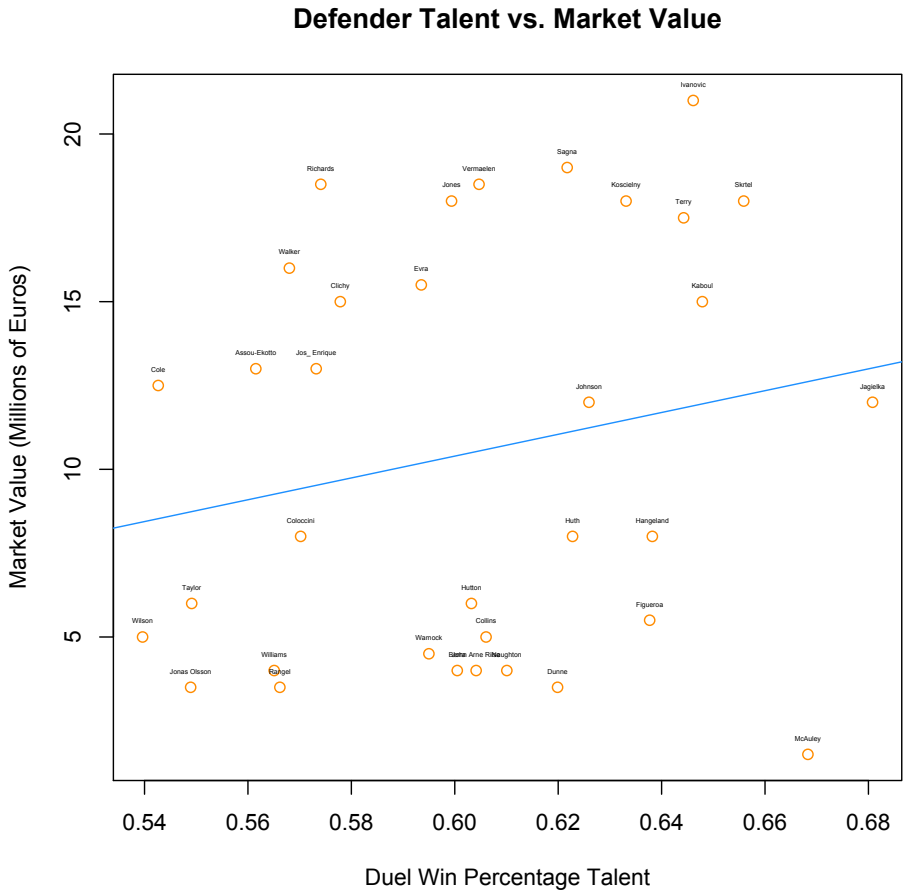
As we can see, the correlation to the value of saves per season is much better when we remove goalkeepers of the "Big Five." Using the distribution above, we obtain a mean market value per save of 212,000 euros for goalkeepers of the "Big Five", and a mean of 47,000 euros for goalkeepers of other teams, with standard deviations of 46,000 euros and 24,000 euros respectively. The estimate makes scoring a goal worth about 15 times saving a goal, which is an indication of the inherent favorability of strikers over goalkeepers in soccer.

[5] Man United, Man City, Liverpool, Arsenal, Chelsea

Although assessing goalkeepers' performance using their save count may seem reasonable, several analysts have actually challenged using this metric. Some analysis quantified the contribution of the quality of shot received to have the most impact on save percentage, and goalkeeper performance[6].

DEFENDERS

When evaluating defenders, multiple metrics could be used to rate a defender as Opta data gives fans tackles, duels, passes, and blocks among other statistics for defenders. However, through experimentation with these various metrics, it was determined that the duel win percentage, calculated by (duels won / total duels) gave the clearest separation in rating the top defenders.



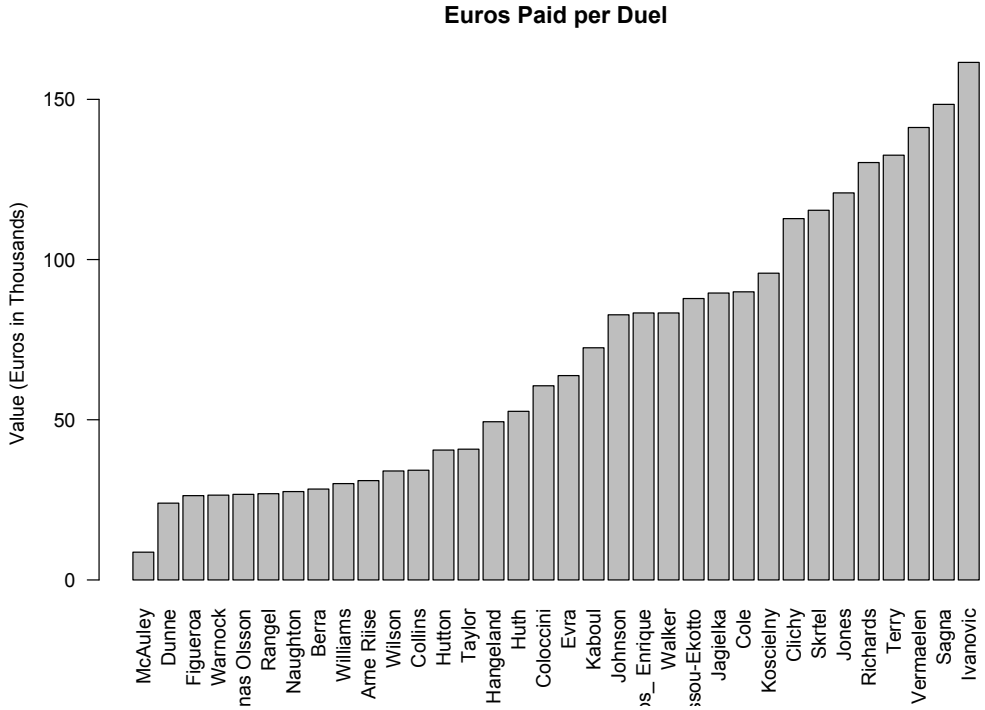
The above graph shows market value versus goals for the most active defenders in the 2011-2012 EPL season, that is, defenders who played in 12 games and won over 125 duels. When looking at defenders in the analysis that are valued in the transfer market at over 15,000,000, one notices that of these players: Ivanovic, Skrtel, Terry, Koscielny, Sagna, Vermaelen, Jones, Richards, Walker, Clichy, Evra, Kaboul, only Younes Kaboul does not come from the traditional "Big Five" clubs in English Football (Arsenal, Liverpool, Manchester City, Manchester United, and Chelsea). This demonstrates that the "Big Five" clubs in English Football are willing to pay large

[6] http://11tegen11.net/2014/02/03/never-judge-a-goal-keeper-by-his-saves/

amounts for the most active defenders in the league and hold who most major clubs deem to be the most valuable defenders.

However, the regression to the mean analysis demonstrates that high-paid defenders are not necessarily the most effective or talented in one on one duels because only five of the ten highest paid defenders fall as part of the top ten talented defenders in the analysis. The bottom-right quadrant of the graph reveals high value defenders because these defenders are, according to our model, active because they have won a high number of duels, talented based on the percentage of duels won, and cost a fraction of the amount in salary as some of the highest paid defenders. The model demonstrates that the following five players are the most high-valued defenders in the 2011-2012 EPL season: McAuley, Jagielka, Figueroa, Hangeland, and Robert Huth. This analysis could be used by mid-table clubs who may want to spend money on a top defender, but are not willing to spend as much money as the "Big Five" clubs do in England.

In professional soccer, particularly with defenders, the media plays a large role in rating players since the position has very few available statistical methods for player evaluations. A recent article in the Liverpool echo entitled, "Phil Jagielka and the stats that prove he's England's most valuable defender," claims that in the 2015-2016 season, Jagielka ranked top two in four of the five key areas of defending (tackles, interceptions, clearances, blocks, and duels). Although Jagielka was a relatively low-profile defender in the 2011-2012 season, using this analysis after the 2011-2012 season would have resulted in signing Phil Jagielka, a player who has developed into one of the best and most consistent defenders in the English Player League.



The value of winning a duel in the 2011-2012 season paid to the most active defenders was 70,000 euros with a standard deviation of 43,000 euros. Again, the large standard deviation relative to the mean dollar amount paid to the most active defenders in English Football demonstrates that again and more so than the previous graph, the high salaries paid by the "Big Five" clubs in English Football creates a huge discrepancy in the wage

data. For example, when looking at the dollar amount paid for duels won, the top 9 players in this metric all come from "Big Five" clubs, again supporting the idea that the "Big Five" clubs are willing to spend more money than other clubs for these important metrics. However, this analysis demonstrates that people are not necessarily paying for the best defenders because relatively two cheap defenders, McAuley and Jagielka, rank highest in our analysis of talent.

CONCLUSIONS

The "Big Five" clubs in English Football are willing to spend and often overspend on the major metrics for forwards, defenders, and goalkeepers that we identified in our study. The highest salaries are paid by these "Big Five" clubs, however, when examining metrics such as talent vs. wage and the dollar amount paid per a certain statistic, one notices that often the highest paid players do not provide the most value for teams. While the "Big Five" clubs in English Football are likely to continue spending inordinate amount on salaries, smaller clubs can use talent and wage analysis to identify talented players who are underpaid and can target these players in the transfer market when looking for high-valued players.

Evaluating player spending in the last 5 years by teams in the English Premier League overwhelmingly supports our hypothesis that these Big Five teams are willing to outspend other clubs on individual statistics. The chart below shows the spending by clubs in English Football. Notice the drop-off after the 5th team. Arsenal (the 5th team in Net Spending) spends more than double what the sixth team in the league has spent on players over the last 5 years.

#	Team	Money Spent on Player Purchases Last 5 Years
1	Manchester City	£472,800,000.00
2	Manchester United	£432,700,000.00
3	Chelsea	£475,959,000.00
4	Liverpool	£354,100,000.00
5	Arsenal	£258,625,000.00
6	West Ham	£120,600,000.00

Our analysis shows that a goal is 18 times more valuable than a save and a goal is 21 times more valuable than a defender duel. As described in the strikers and goalkeepers section of the paper, action count correlates better with market value than skill level for both goal-scoring and goal-saving. However, with winning duels we see very low correlation (0.2) with market value versus player talent and an even lower correlation when considering action count. This makes sense when considering the description of the positions. While goalkeepers and forwards are primarily tasked with saving and scoring goals respectively, defenders are asked to win duels, tackle, complete passes, and for outside backs, get forward into the attack. Since we are considering outside backs and center backs within the defenders category, the low correlation is not surprising due to the variance and range of task required by the position.