A Proposed Decision Rule for the Timing of Soccer Substitutions

Bret R. Myers, *Villanova University*

A Proposed Decision Rule for the Timing of Soccer Substitutions

Bret R. Myers

Abstract

Managers in soccer face a critical in-game decision concerning player substitutions. While past research has investigated the overall effect of various strategy changes during the course of a match, no studies have focused solely on the crucial element of timing. This paper uses the data mining technique of decision trees to develop a decision rule to guide managers on when to make each of their three substitutes in a match. The proposed decision rule demonstrates between 38% and 47% effectiveness when followed versus 17% and 24% effectiveness when not followed based on over 1200 observations collected from some of the world's top professional leagues and competitions.

KEYWORDS: analytics, data mining, soccer, substitutions, football

1. Introduction

Contrary to other sports, soccer managers have limited opportunities to directly impact the course of the match. Once the whistle blows to begin play, most of the control is relinquished to the players on the field. With no timeouts and the only guaranteed stoppage in play occurring at half-time, managers for the most part take a back seat while the players on the field author the majority of in-game tactical decisions. However, managers do maintain ownership in the crucial decision of player substitutions. FIFA, the international governing body of soccer, restricts the number of substitutions to three and does not allow re-entry of the substituted player. Accordingly, most professional leagues and tournaments around the world subscribe to these rules. The stakes are the highest in this domain which includes the World Cup and top professional leagues such as the English Premier League, La Liga in Spain, Serie A in Italy, the Bundesliga in Germany, and MLS in the United States. In many instances, substitutions can be the determining factors that make or break a team's performance. Under such constraints, substitutions become scarce resources that managers must allocate properly. Therefore, the fundamental goal of this paper is to attempt to develop an effective substitution strategy which enhances the probability of achieving desired results.

So what is an appropriate substitution strategy? Are there critical moments when a substitute may be a better alternative to a starter who is either fatigued or exhibiting poor play? Does the value of a substitution vary according to whether or not a team is ahead, tied, or behind in a match? These are the questions that this paper attempts to answer through analytical methods applied to historical data.

Professional soccer organizations and coaches have traditionally shied away from relying on mathematics and statistics to shape coaching strategies. However, there are stories of success in how teams being able to use analytical methods to gain a competitive advantage. Kuper and Symanski (2009) devote a chapter in their widely popular book *Soccernomics* to penalty kick strategies. The authors include anecdotes of teams that built a record of the tendencies of penalty kick takers in order to use it to their advantage in situations like a penalty shootout. Two examples mentioned were the German National Team in their shootout against Argentina in World Cup 2006, and Chelsea in their shootout against Manchester United in 2008.

AC Milan has been cited as an organization that uses analytics in the arena of injury prevention and player sustainability (Aziza 2010). The club won the 2007 Champions League title while fielding a veteran squad with an average age of 31. Recently in their 2009-2010 Champions League campaign, the club fielded a backline with an average age of 32.5 (Anderson 2010). The ability to keep

veteran players healthy and playing at an elite level is largely attributed to their extensive research combining science, technology, and statistical methods.

Sam Allardyce in his tenure as manager at English Premier League club Bolton was an early adopter of extensive video analysis to produce useful statistics to improve team performance (Davenport 2007). Over the past couple of years, video analytics has become increasingly popular among professional soccer organizations. A California-based video analytics company named Match Analysis is widely used in both the MLS and WPS, and also has expanded its customer base to include teams from other professional leagues around the world, various national teams, and several domestic college and high school teams (Match Analysis). The European-based company Opta has grown rapidly in the past few years as it is now a popular choice among the top professional leagues around the world (Opta Sports). Through partnerships with these types of advanced analytics companies, soccer organizations have the capability to analyze mountains of meaningful data in order to gain a competitive advantage.

Thus, soccer organizations have demonstrated a commitment to invest in methods to break down the game of soccer into objective measures that can better explain a team's performance. A major goal of this paper is to provide fact-based information on substitution strategy that can help managers in practice.

2. Research in Soccer Substitution Strategies

The problem of how to substitute in soccer has garnered the attention of researchers in the past. Hirotsu and Wright (2002) use dynamic programming to find optimal time to "change tactics" based on data from the English premier league. The authors incorporate variables such as whether or not the team is playing at home or away, the time remaining during the match, and the tactical formation on the field. A tactical change not only can be a substitution but also a change of formation imposed by the manager. Although there are some interesting findings concerning what teams should do based on their current state of goal differential during the match, there are no clear recommendations for managers on how to implement the proposed models.

Hirotsu et al. (2009) use game theory to model the substitution problem as a non-zero sum game. They suggest moments when it is best to change formation based on whether or not the team is home away, the time remaining in the match, the current goal differential, and the formation of the opponent. Similar to the previous paper discussed, it is not clear how managers can use the authors' findings in practice.

Corral et al. (2007) analyze substitutions based on the 2004-2005 Spanish league season. The authors determine that the timing of the first sub is based on goal differential and that there is evidence that teams that are behind sub earlier than those that are tied or ahead. In addition, it was found that defensive subs

occurred later in the match while offensive subs occurred earlier. While the study is effective in summarizing the tendencies for the Spanish League, the results cannot necessarily be universalized for all leagues. At the same time, the research fails to provide managers with a specific guide on how to substitute.

This paper will focus on the criticality of the timing of each of the three substitutions during a soccer match. This is an important component of the substitution decision which can be often overlooked. Managers have a tendency to be more reactive than proactive with their substitutions. Although players are conditioned to play a full ninety minutes, their performance levels may drop later in matches. If managers take the approach of waiting for signs to indicate that a starter on the field needs to be replaced, it is likely that the critical moment to substitute that particular player has already passed. Davenport (2007) reports that analysts for the Boston Red Sox determined in the 2003 MLB season that the performance level of ace picture Pedro Martinez dropped after about 7 innings or 105 pitches. Accordingly, it was determined that in pitches 91-105, opposing batters hit .231, while in pitches 106-120, the batting average by opposing batters increased to .370. In a decisive game of the ALCS against the New York Yankees that season, Martinez's pitch count reached the designated threshold level, but manager Grady Little ignored the provided statistics and let Martinez continue to pitch. The outcome was a Yankees hit fest in the 8th innings, a Red Sox exit from the playoffs, and the end of the Grady Little campaign in as manager in Boston. The following season, the Red Sox went on to win the World Series, most likely with the virtue of increased attention and application of analytical findings.

The Pedro Martinez pitch count case is an example of how a manager overvalued a starter and undervalued the role of a reliever pitcher. Is there a tendency in soccer for managers to do the same? This paper will provide an answer to this question. Although the decisions of which players to substitute and the choice of formation are also important, it is not within the intentions of this paper to provide the answers to these questions.

3. Data Collection

In order to better understand substitutions, data has been collected representing 485 observations of a particular team's substitution patterns. The variables tracked are goal differential before and after each substitution, whether or not a team was home or away, and the timing of each substitution. Observations of 155, 172, and 158 are drawn respectively from the English Premier League, Serie A in Italy, and La Liga in Spain using data provided by ESPN Soccernet. Team formation is not taken into consideration since the variable is difficult to tract and can also be considered subjective. At the same time, a team could start in one formation, then evolve into at least one other during the course of the match.

Past research has also focused on whether or not a substitution is considered attacking or defending. This paper does not make a distinction between these two roles. Although players of the game of soccer hold a specific position in a team's formation, all players on the field have a role in both the attack and defense. In addition, it would be difficult to classify whether or not a substitution was either attacking or defended by looking at a box score and not knowing enough about the players on the team.

4. Exploratory Analysis

The next step is to perform exploratory analysis on the data which include summary measures and hypothesis testing. The variables of interest in particular are substitution times and the number of subs. Table 1 provides the summary statistics on substitution times and Figure 1 is a histogram of substitution times.

Table 1: Summary Statistics on Substitutions

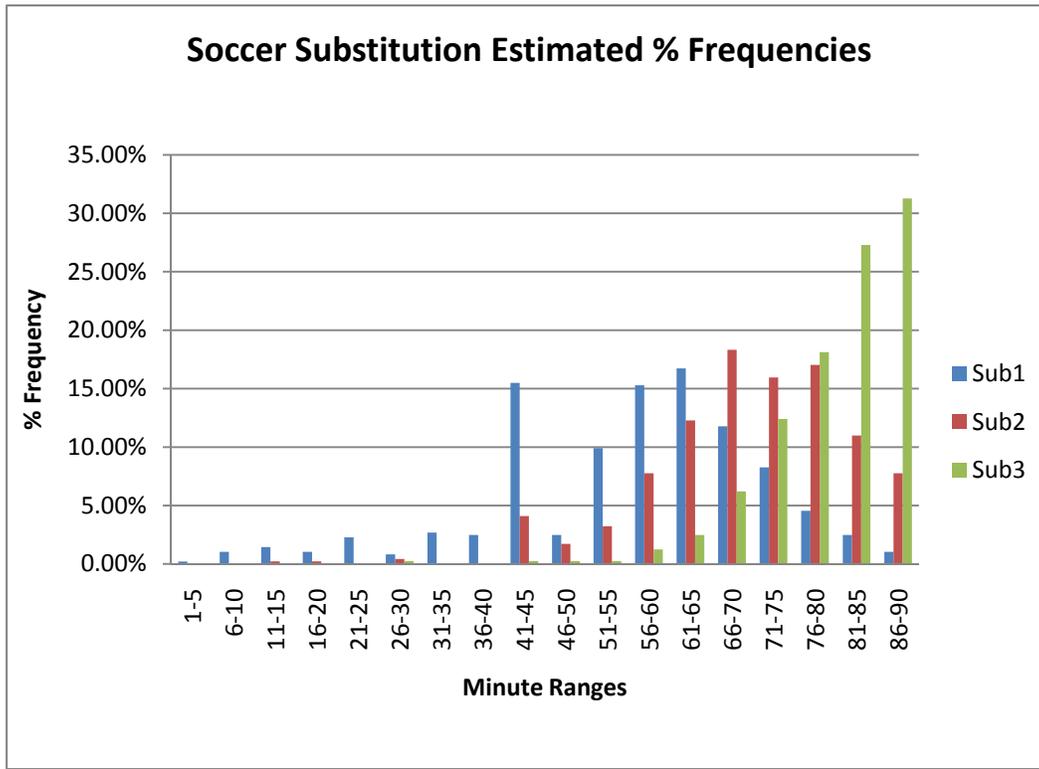| Substitution Summary Stats			
Sub	**1**	**2**	**3**
Mean	56.38	70.23	80.58
Median	59	71	82
Mode	45	70	90
Stdev	15.6	11.75	8.22
N	484	464	403

Figure 1: Histogram of Substitution Times

Beginning with the first sub, the average is approximately around the 57th minute while the median is around the 60th minute. The distribution is skewed by first half substitutions, the majority of which are most likely injury related. The mode is at 45 minutes which is considered half-time. This makes sense since it is a stoppage in play and the manager is afforded the opportunity to address the team. The distribution of the second sub is more symmetric around the 71st minute, a little past the midpoint of the 2nd half. The distribution of the third sub is slightly left skewed, mainly since the distribution is capped at 90 minutes. It is worth noting that the mode is also 90 minutes. This is most likely true because of injury time that occurs after the 90th minute. Any sub in injury time is technically charted as the 90th minute. With that said, the mean and median of the 81st minute and 82nd minute are not far apart.

The next variable of interest is the proportion of teams that use a particular number of subs. A bar chart is provided in Figure 2.
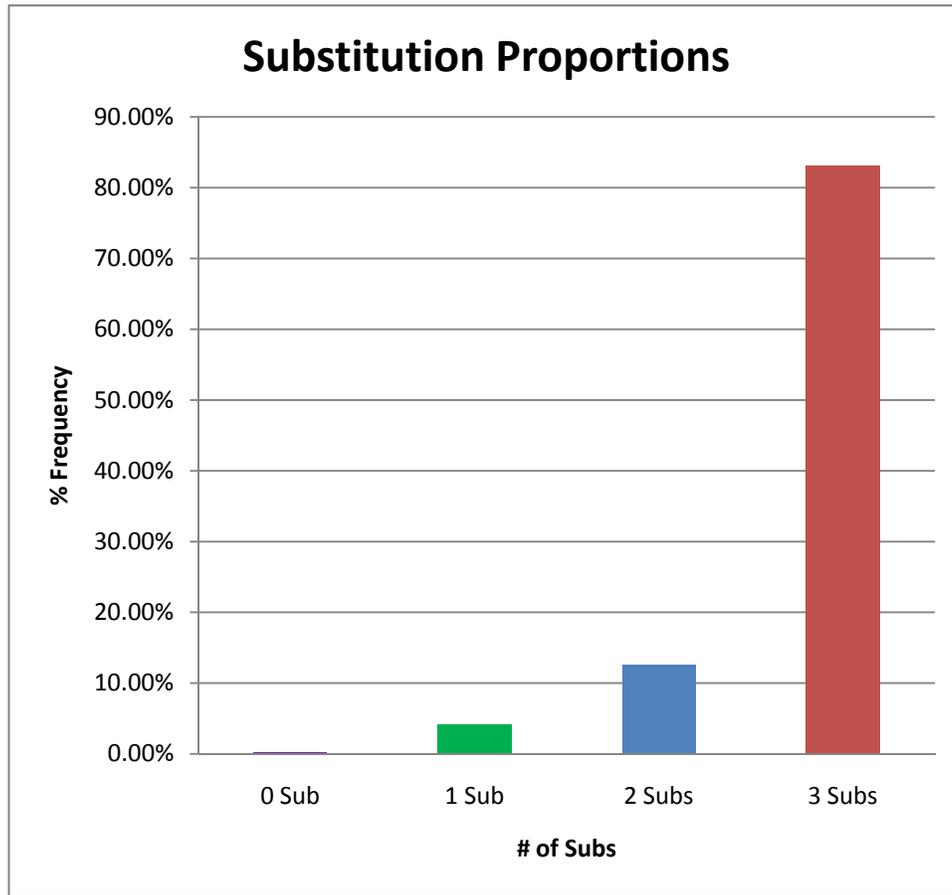
Figure 2: The proportion of teams using 0,1,2, or 3 subs

It is clear that the majority of teams use all three substitutions. This is indicative of the fact that most teams value the contribution of reserve coming off the bench. The performances of using 3 subs vs. 2 subs or less are summarized in Table 2. The results are classified based on whether a team is down, tied, or up prior to their substitutions

Table 2: Performance according to number of subs made

Win/Tie % based on scenario and # of Subs				
Scenario	3 Subs	N	<2 subs	N
Down	24.20%	123	9.52%	21
Tied	72.85%	151	61.11%	18

Although the sample size is small for the case of two or fewer substitutions, teams that used all three subs seem to outperform those that don't when down or tied in a match. Again, this supports the idea that substitutes can be of more value than existing players of the field in the event that a team is tied or down.

In order to gain further information from the data, three hypothesis tests are conducted concerning substitutions and the variables considered in the study. For each of these hypothesis tests, it is assumed that substitutions are independent between games. There is also an underlying assumption that minimal correlation exists between substitution times by a team within a game. The purpose of these tests is to help determine factors that affect substitution times within a game.

Hypothesis 1: The timing of each sub varies according to a team's score position (losing, tied, or ahead).
Result: There is a statistically significant difference for each of the three subs

Table 3 below provides the results of the one-way ANOVA procedure using for each of the score positions respectively:

Table 3: The effect of score position on substitution times

The Effect of Current Score Position			
	Sub1	Sub2	Sub3
Avg|Down	53.69	64.65	76.87
Avg|Tied	51.07	68.21	80.19
Avg|Up	59.61	73.44	84.06
p-value	9.09E-06	1.28E-10	4.10E-14

The results indicate that managers tend to hold onto subs later when the team is ahead, but make substitutions earlier when either tied or behind. The intuition behind these results is that managers value starters more when they are able to produce a lead for the team, so they are willing to let them continue to play longer than jeopardizing the flow of good play with a player change. However, managers still need to be mindful of player fatigue, so the substitution will not occur too late. When a team is tied or ahead, managers become more risk-seeking by plugging in a substitute earlier to try to inject new energy into a team that is faced with the task of trying to improve the current score state of the team.

Hypothesis 2: The timing of substitutions differ according to whether or not a team plays at home
Result: No significant difference on subs 1 and 2, but slight difference on sub 3.

Table 4 below provides the result of the two sample t-test for the difference in mean away vs. home substitution time.

Table 4: The effect of Home/Away on Substitution times

The effect of Home/Away			
Sub#	Away	Home	p-value
Sub1	53.8	55.27	0.341
Sub2	68.09	69.61	0.182
Sub3	79.79	81.33	0.06

One possible explanation for the slight significant difference in the third substitution is it has been shown that managers substitute earlier when a team is down, and the away team has an increased likelihood of being down in a match. Since there is less variation in the timing of the third sub, then difference ended up being slightly significant.

Hypothesis 3: The timing of substitutions differs according to league type.
Result: There is a significant difference with the first sub, but not the second and third substitutes.

Table 5 reports the results of the ANOVA procedure for each substitution.

Table 5: The Effect of League Type on Substitutions

The Effect of League Type			
League	1st Sub	2nd Sub	3rd Sub
English	57.77	69.58	80.08
Italian	52.06	68.27	80.46
Spanish	55.07	69.64	81.87
p-value	0.011	0.657	0.275

It is interesting to see a difference in league type with the first sub. The differences may be rooted in cultural values. The Italian league and English were significantly different from each other with the first sub, possibly suggesting that Italian managers are less patient with their starting eleven than English managers.

5. Data Mining

The next step in the analysis is to determine if a substitution strategy can be developed through the use of the data mining technique of decision trees. The methodology will be used to find optimal splits in substitution times which lead to enhanced probability of success. One major advantage of decision trees is that they are easy to interpret. One of the primary objectives in this paper is to provide managers with practical rules that are easy to follow. Using the technique of decision trees can support this objective.

Based on the results from the exploratory analysis, it makes sense to try to develop a rule based on the current score state of a team. If a team is behind in a match, the objective is to try to come back. An improvement will be defined as an increase from the current score state prior to a substitution. A binary variable can be used to reflect whether or not a team was able to improve the score state prior to a sub. A value of "1" denotes an improvement and a value of "0" denotes no improvement.

When a team is tied, the target variable is goal differential. Positive goal differential indicates that a team was able to win as a result of a substitution; zero goal differential indicates that they preserved a tie, and negative goal differential indicates that team ends with a loss after a substitution. On the other hand, when a team is ahead, the target variable is binary. A "1" denotes that the win was preserved after a substitution and a "0" indicates that a tie or loss occurred after the substitution.

SAS® Enterprise Miner is used to run the decision tree algorithm. The original 485 observations are used as training data. The results indicate slightly significant splits in the data when a team is behind. The splits are presented in Table 6.

9

Table 6: Results of Decision Tree Algorithm

Decision Tree Results If down			
	sub1	sub2	sub3
avg target	0.31	0.24	0.16
Split	57.5	72.5	78.5
avg target below split	0.41	0.3	0.24
avg target above split	0.18	0.06	0.07
p-value	0.052	0.057	0.134

Beginning with the first sub, the results indicate that teams that were down prior to the first sub improved 41% of the time when substituting before the 58th minute, regardless of the 2nd and 3rd subs, while subbing at the 58th minute or later only led to an improvement 18% of the time. With the 2nd sub, a significant split occurred at between then 72nd and 73rd minute. Teams behind when needing to make a 2nd sub achieved improvement 30% of the time when substituting prior to the 73rd minute but only achieved improvement 6% of the time when not substituting prior to the 73rd minute. Finally, with regard to the 3rd substitute, there was only weak significance in a split. Teams that substituted prior to the 79th minute had a 24% improvement rate vs. a 7% improvement rate for teams that waiting until the 79th minute or later.

In all these cases of substituting when a team is behind, there was not an identified minimum point on when each sub should take place. Most substitutions in soccer take place at either half time or later. The rare cases of first half substitutions are generally for the purpose of injury or extreme adjustments needed to be made following a player ejection. Thus, it is generally recommended to wait until half time to make a substitution, although there is no statistical evidence to support this claim.

The decision tree procedure did not produce any significant splits for when a team is tied or ahead. This is evidence to suggest that the timing of a sub is not as critical as when a team in behind in the match. Thus, managers should place more weight on other criteria when making substitution decisions under these settings.

6. Proposed Decision Rule

Based on the results of the data mining study, the splits can be combined to form the decision rule explained in Figure 3.

Proposed Decision Rule
- If down
  - Make 1st sub prior to 58th minute
  - Make 2nd sub prior to 73rd minute
  - Make 3rd sub prior to 79th minute
- If tied or ahead
  - Make subs at will

*The rule does not apply under the following conditions:
- Forced substitutions in the first half due to injury
- Red cards committed by either team
- Matches that have overtime if the game is tied after regulation

Figure 3: Guidelines of the Proposed Decision Rule

As the game approaches the first critical point of the 58th minute, a coach should make at least the first substitute if behind. As the game approaches the next critical point of the 73rd minute, if still behind, a coach should make at least the 2nd substitute. If the team is able to equalize or go ahead once the critical point is reached, then it is allowable for the 2nd substitute to be withheld. However, if the team returns to a state of being behind prior to the last critical point of the 79th minute, then a coach should use both the 2nd and 3rd substitution prior to the 79th minute. If a team that was previously tied or ahead falls behind after the 80th minute, there is no specific recommendation on how a coach should use the remaining substitutes if still available.

An instance is defined as an opportunity during a match where a particular team falls behind and can apply the decision rule prior to at least one of the three critical points (58th, 73rd, or 79th minute). It is possible have two instances occur during one game. If a team that is originally behind goes ahead of the opposing team, then the opposing team could be put into a position to apply the decision rule. It is not possible to have more than two instances, but it is possible for no instances to occur.

11

Since there were no significant results found for when a team is tied or ahead, the decision rule is believed to unaffected by the substitution pattern exhibited by the opposing team. The opposing team would be tied or ahead, therefore, there is no clear strategy that could be employed to combat the effects from the team from behind following the decision rule.

In order to validate the rule, a collection of observations were selected from the English Premier League, La Liga in Spain, Serie A in Italy, the German Bundesliga, MLS in the USA, and the World Cup 2010. A total of 1284 instances were analyzed where managers had a clear opportunity to apply the proposed decision rule. All instances were analyzed for the 2010 World Cup, 2009-2010 English Premier League, and 2010 MLS Season, while a random sample was collected from the both the 2009-2010 and 2010-2011 seasons of La Liga in Spain, Serie A in Italy, and the Bundesliga in German. There is some duplication in the training data and test data, but enough to jeopardize the validity of the study. The Bundesliga, MLS, and World Cup were not part of the training data. In addition, not all observations in the original training set represented instances where the decision rule could be applied.

A success from following the rule indicates that the team was able to improve their goal differential from the first moment a team substitutes in adherence to the guidelines of the rule to the end of the match. Matches that involved substitutions due to injuries in the first half, or red cards by either team during influential moments are discarded from consideration. A failure from following the rule indicates that the team could not improve goal differential. Teams that do not follow the rule elect to hold at least one substitute after at least one of the three corresponding critical moments. Success is achieved if goal differential is improved from the beginning of the earliest critical moment bypassed and the end of the game.

7. Performance of Decision Rule

The overall results on the test of the performance rule are reported in Table 7.

Table 7: Performance of Decision Rule in the cumulative test sample

Measure	Decision Rule	No Decision Rule	Total
Number of successes	186	173	359
Number of failures	254	670	924
Instances	440	843	1283
Success Rate	42.27%	20.52%	27.98%
Participation Rate	34.29%	65.71%	100%

Teams that followed the decision rule had a success rate of 42.27% versus 20.52%. Based on 95% confidence limits, the success rate for teams following the decision rule is between 38% and 47% while the success rate of teams not subscribing to the rule is between 17% and 24%. This shows that there is a statistically significant difference in the performance of two substitution policies, and team can roughly double their chance of improvement by subscribing to the proposed decision rule. The overall participation in the rule across all instances sampled is only 34.29%. Therefore, the managers in the majority of applicable instances are not taking advantage of a strategy that demonstrates a better success rate.

An exhaustive study was conducted from the 2009-2010 seasons in Italy, England, Spain, and Germany, as well as MLS 2010 and World Cup 2010. A comparison of how well the decision rule performance across the various competitions is presented in Table 8.

Table 8: The performance of the decision rule by league/competition

League/Tournament	Success w/Rule	Success w/o Rule	Decision Rule participation %
2009-2010 Serie A (Italy)	52.50%	17.83%	38.28%
2010 World Cup	45.45%	14.29%	44.00%
2009-2010 Bundesliga (Germany)	43.28%	25.00%	27.54%
2009-2010 Premier League (England)	40.00%	21.59%	23.81%
2010 MLS (USA)	36.36%	19.35%	41.51%
2010 La Liga (Spain)	32.35%	21.13%	32.38%

In all competitions, the decision rule lead to statistically better success percentage. The level of success varied, with a maximum rate of 52.50% exhibited in Serie A and a minimum rate of 32.35% exhibited in La Liga. The discrepancy between the two leagues here demonstrates a statistically significant difference in success rate. One could speculate that the characteristics of the Italian game vs. the Spanish game may help to explain this difference. The levels of participation also varied. Albeit a small sample size, there was 44% participation in the World Cup, and a league high 41.51% participation found in MLS 2010. The English Premier League's 23.81% participation is the lowest among the competition's analyzed.

More detailed results from each of the five leagues are included in Appendices 1-5. On a team-by-team basis, there were some clear winners and losers as far as the decision rule goes. Jose Mourinho, recent recipient of the 2010 World Manager of the year award, led Inter to the Serie A and Champions League titles. During 7 of the 38 games played during the 2009-2010 Serie A season, Mourinho's squad was presented with an opportunity to subscribe to the proposed decision rule. In 6 out of the 7 instances (85.71%), Mourinho followed the rule, and was successful in 5 out of the 6 times (83.33%). The one time that he did not follow the rule led to a failure. This performance can be compared to rival club Juventus, managed by the much less renowned Ciro Ferrara. Juvenus had 11 opportunties to follow the decision rule and did so only twice (18.18%). Juventus was successful on one of the two occasions in following the rule (50%). On the 9 occasions of not following the rule, Juventus was only successful once (11.11%). Nevertheless Juventus finished in the dreaded position of 7th which precluded them from European competition, and the manager was sacked at the end of the season.

A similar tail was demonstrated in Germany when comparing Bundesliga champions and UEFA champions league runners-up Bayern Munich to 5th place finisher Borussia Dortmund. In a 34 game season, highly regarded Bayern Munich manager Louis Van Gaal followed the decision rule 5 out of 8 opportunities presented (62.50%) and was successful 4 out of the 5 times (80%). Out of the three instances when the rule was not followed, Bayern Munich only enjoyed success 1 out of 3 times. On the other hand, Borussia Dortmund manager Juergen Klopp only participated in the rule 2 out of 9 opportunities presented (22.22%), yet they were successful on both occasions. Out of the 7 instances where the rule was not followed, Dortmund went 0 for 7. Unlike Ferrara of Juventus, Klopp of Dortmund still kept his job entering the 2010-2011 season.

What stood out most about the 2009-2010 English Premier League season were the missed opportunities with the 23.81% participation rate. Two of the most glaring culprits were Liverpool and Portsmouth, who both had disappointing finishes relatively speaking. Liverpool, a usually fixture in England's top 4,

finished the dreaded 7th place like their Italian counterparts Juventus. Manager Rafa Benitez only seized the opportunity to follow the decision rule 1 out of 9 times (11.11%). There were successful on the one occasion, but failed on all eight of the instances where the rule was not followed. Needless to say, Benitez was let go at the end of the season. The other poor display regarding the decision rule was shown by last place finishers Portsmouth. Manager Avram Grant faced a league high 21 instances to apply the decision rule. He only participated 4 out of the 21 times (19.05%), yet was successful in two of the occasions (50%). In the 17 occasions where the rule was not applied, Portsmouth only found success 3 out of 17 times (17.65%). As the club exited the Premier League that season, so did Manager Grant from his post.

The MLS season did not produce any clear winners or losers. Although both teams failed to reach the playoffs, Houston and Toronto both had high participation and success rates with the decision rule, a tribute to managers Dominic Kinnear and Preki respectively. Last place finisher DC United suffered some misfortunate as they were only about to achieve success 1 out of the 14 applicable opportunities. It was a rocky season for DC United in that Ben Olson replaced Curt Onalfo as manager during the middle of the season. DC United had high participation in the decision rule (62.29%), and by following this strategy, they achieved there only success.

Lastly, La Liga demonstrated the lowest level of success (32.28%) as the participation rate was also relatively low at 32.35%. The most striking example was that of 19th place finishers Tenerife. Manager Gonzalo Arcanada was presented with 16 opportunities to follow the decision rule, and neglected to do so in every instance. Tenerife only achieved success on 2 of these occasions (12.50%). Tenerife was relegated down to second division, and Arcanada was dismissed for his position in September of 2010.

8. Conclusion

Overall, this paper reveals a valuable and practical decision rule for managers to follow regarding their three substitutions. The qualifying conditions for application of the rule are that the team must be behind in the match approaching any one of the three suggested critical points of the 58th, 73rd, and 79th minutes, and no red cards or forced first half injuries have occurred. In subscribing to the proposed policy, teams across the board are believed to have a success rate of 38% to 47% vs. 17% to 24% in decreasing goal differential when behind. The rule was only followed 34.29% of the time based on the 1,284 instances analyzed.

While the proposed strategy tests well, it is not guaranteed to be optimal in every case. Every soccer game is essentially unique; therefore, it is difficult to prescribe an exact rule according based on all the possible variables that can affect substitution times. The value of the decision rule proposed in this study is

that it is a general solution that probabilistically dominates its complement rule. Future research will be devoted to the incorporation of more factors that affect substitutions, which may lead to better results than reported in this paper.

Overall, the results from this study suggest that managers tend to overvalue starters on the field and undervalue the role of substitutes. When a team is behind in the match, players on the field can easily become both mentally and physically fatigued. A fresh substitute can inject some needed energy into the squad and provide a spark that is less likely to occur from existing players on the field. In quality leagues like the English Premier League, Spanish La Liga, the German Bundesliga, and MLS of the USA, managers typically have quality reserves on the bench who are capable of being starters. Therefore, managers should not be as hesitant to make substitutions since they are likely to have at least three capable players on the bench.

In closing, this paper demonstrates how analytics can be used to provide managers with useful information to enhancing their coaching strategies. In an era of increased sponsorship of technology and analytical methods, managers must find the right balance of reliance on such techniques and natural gut feeling and intuition. Ultimately, a soccer team should not be run by a computer as managers have the unique ability to weigh in on the human elements of the game that are not easily quantifiable. At the same time, when being the recipient of knowledge generated by a sound analytical study, the soccer manager is clearly better off than being ignorant in the matter. Concerning substitutions, it is valuable to know that there is a strategy that has shown general success for all teams. This paper offers managers the opportunity to possess new information on substitutions that can be easily integrated into their overall coaching strategy.

Appendix 1

2009-2010 Italian Serie A

Performance of Decision Rule

League Position	Team	# of instances	decision rule participation	decision rule success	non-decision rule success
1	Inter	7	85.71%	83.33%	0.00%
2	Roma	4	25.00%	100.00%	66.67%
3	AC Milan	4	25.00%	0.00%	66.67%
4	Sampdoria	6	66.67%	50.00%	0.00%
5	Palermo	13	30.77%	50.00%	44.44%
6	Napoli	8	25.00%	50.00%	33.33%
7	Juventus	11	18.18%	50.00%	11.11%
8	Parma	9	33.33%	33.33%	0.00%
9	Genoa	6	50.00%	33.33%	0.00%
10	Bari	13	46.15%	66.67%	42.86%
11	Fiorentina	9	33.33%	100.00%	16.67%
12	Lazio	10	40.00%	75.00%	16.67%
13	Catania	6	33.33%	100.00%	25.00%
14	Chievo	13	23.08%	33.33%	30.00%
15	Udinesse	15	33.33%	80.00%	10.00%
16	Cagliari	17	29.41%	60.00%	16.67%
17	Bologna	11	27.27%	66.67%	0.00%
18	Atalanta	15	53.33%	12.50%	0.00%
19	Siena	16	43.75%	42.86%	0.00%
20	Livorno	16	50.00%	25.00%	0.00%
	Total	**209**	**38.28%**	**52.50%**	**17.83%**

17

Appendix 2

2009-2010 German Bundesliga

Performance of Decision Rule

League Position	Team	# of instances	decision rule participation	decision rule success	non-decision rule success
1	Bayern Munich	8	62.50%	80.00%	33.33%
2	Schalke	5	40.00%	50.00%	0.00%
3	Werder Bremen	11	27.27%	50.00%	71.43%
4	Bayer Leverkusen	7	71.43%	60.00%	100.00%
5	Dortmund	9	22.22%	100.00%	0.00%
6	Stuttgart	12	50.00%	42.86%	0.00%
7	Hamburg	7	0.00%	N/A	28.57%
8	Wolfsburg	13	30.77%	50.00%	11.11%
9	Mainz	10	40.00%	0.00%	50.00%
10	Eintracht Frankfurt	12	16.67%	50.00%	50.00%
11	TSG Hoffenheim	13	23.08%	50.00%	22.22%
12	M'gladbach	14	7.14%	0.00%	15.38%
13	Cologne	9	0.00%	N/A	33.33%
14	SC Freiburg	15	20.00%	0.00%	18.18%
15	Hannover 96	15	26.67%	40.00%	0.00%
16	Nurnberg	19	36.84%	50.00%	0.00%
17	Bochum	14	35.71%	25.00%	0.00%
18	Hertha	14	7.14%	50.00%	0.00%
	Total	**207**	**27.54%**	**43.28%**	**25.00%**

Appendix 3

2009-2010 English Premier League

Performance of Decision Rule

League Position	Team	# of instances	decision rule participation	decision rule success	non-decision rule success
1	Chelsea	3	33.33%	0.00%	50.00%
2	Man U	10	20.00%	50.00%	25.00%
3	Arsenal	8	50.00%	25.00%	0.00%
4	Tottenham	9	33.33%	0.00%	16.67%
5	Man City	4	50.00%	100.00%	0.00%
6	Aston Villa	9	22.22%	100.00%	42.86%
7	Liverpool	9	11.11%	100.00%	0.00%
8	Everton	8	0.00%	N/A	50.00%
9	Birmingham	12	8.33%	100.00%	27.27%
10	Blackburn	12	58.33%	28.57%	20.00%
11	Stoke	8	12.50%	0.00%	14.29%
12	Fulham	16	18.75%	33.33%	46.15%
13	Sunderland	13	7.69%	0.00%	16.67%
14	Bolton	14	7.14%	0.00%	7.69%
15	Wolves	15	26.67%	25.00%	18.18%
16	Wigan	16	12.50%	0.00%	21.43%
17	West Ham	12	41.67%	40.00%	14.29%
18	Burnley	17	35.29%	66.67%	27.27%
19	Hull	15	33.33%	40.00%	10.00%
20	Portsmouth	21	19.05%	50.00%	17.65%
	Total	**231**	**23.81%**	**40.00%**	**21.59%**

19

Appendix 4

2010 MLS

Performance of Decision Rule

League Position (by points)	Team	# of instances	decision rule participation	decision rule success	non-decision rule success
1	Los Angeles	8	50.00%	25.00%	25.00%
2	Real Salt Lake	6	33.33%	100.00%	25.00%
3	New York	6	0.00%	N/A	16.67%
4	Columbus	9	22.22%	50.00%	42.86%
5	Dallas	7	28.57%	100.00%	40.00%
6	Seattle	9	33.33%	0.00%	0.00%
7	Colorado	8	25.00%	50.00%	0.00%
8	San Jose	9	44.44%	25.00%	0.00%
9	Kansas City	12	41.67%	40.00%	28.57%
10	Chicago	9	44.44%	50.00%	0.00%
11	Toronto	12	50.00%	50.00%	0.00%
12	Houston	9	55.56%	60.00%	25.00%
13	New England	13	23.08%	0.00%	20.00%
14	Philadelphia	10	60.00%	33.33%	50.00%
15	Chivas	18	50.00%	22.22%	33.33%
16	DC United	14	64.29%	11.11%	0.00%
	Total	**159**	**41.51%**	**36.36%**	**19.35%**

Appendix 5

2009-2010 Spanish La Liga

Performance of Decision Rule

League Position	Team	# of instances	decision rule participation	decision rule success	non-decision rule success
1	Barcelona	0	N/A	N/A	N/A
2	Real Madrid	4	25.00%	100.00%	0.00%
3	Valencia	6	16.67%	0.00%	20.00%
4	Sevilla	7	42.86%	66.67%	50.00%
5	Mallorca	13	23.08%	33.33%	30.00%
6	Getafe	9	33.33%	0.00%	16.67%
7	Villareal	9	55.56%	40.00%	25.00%
8	Athletic	11	54.55%	33.33%	20.00%
9	Atletico	7	14.29%	0.00%	0.00%
10	Deportivo	14	28.57%	25.00%	10.00%
11	Espanyol	11	45.45%	40.00%	0.00%
12	Osasuna	10	0.00%	N/A	10.00%
13	Almeria	14	14.29%	50.00%	25.00%
14	Zaragoza	15	20.00%	33.33%	16.67%
15	Sporting	13	23.08%	33.33%	10.00%
16	Racing	16	18.75%	66.67%	30.77%
17	Malaga	9	44.44%	50.00%	60.00%
18	Valladolid	9	44.44%	25.00%	20.00%
19	Tenerife	16	0.00%	N/A	12.50%
20	Xerez	17	29.41%	60.00%	25.00%
	Total	**210**	**32.38%**	**32.35%**	**21.13%**

References

Anderson, Michael. "Football Special Report: AC Milan Must Sign a Centre Back in January" Better.com. 12 Dec. 2010. 15 Feb. 2011. http://blogs.bettor.com/Football-Special-Report-AC-Milan-must-sign-a-centre-back-in-January-a47894

Aziza, Bruno. "The Future of Soccer is in its Analytics". Smart Data Collective. 6 Jun. 2010. 10 Feb. 2010. http://smartdatacollective.com/Home/27719

Corral, J., C.P. Barros, and J. Prieto-Rodriguez. "The Determinants of Player Substitutions: A Survival Analysis of the Spanish Soccer League." The Journal of Sports Economics.

Davenport, Thomas H. and J.G. Harris. *Competing on Analytics: The New Science of Winning*. Boston, MA: Harvard Business School Press, 2007.

ESPN Soccernet. http://soccernet.espn.go.com/?cc=5901

Hirotsu, N. and M. Wright. "Using a Markov Process Model of an Association Football Match to Determine the Optimal Timing of Substitution and Tactical Decisions" Journal of the Operations Research Society 53.1: 2002 pp. 88-96.

Hirotsu, N., M. Ito, C. Miyaji, and K. Hamano. "Modeling Tactical Changes in Football as a Non-Zero-Sum Game." Journal of Quantitative Analysis in Sport 5.3: 2009

Kuper, Simon and S. Syzmanki. *Soccernomics*. New York: Nation Books, 2009.

Match Analysis. 10 Feb, 2011. http://www.matchanalysis.com/

Opta Sports. 14 Jul, 2011. http://www.optasports.com/