
Munich Personal RePEc Archive

Do some teams consistently underperform against same opponents? - The curious case of the Croatian football league

Globan, Tomislav and Jägers, Ed

Faculty of Economics and Business, University of Zagreb,
DAT.Mobility, Deventer, the Netherlands

20 November 2016

Online at https://mpra.ub.uni-muenchen.de/75082/
MPRA Paper No. 75082, posted 21 Nov 2016 09:05 UTC

Do some teams consistently underperform against same opponents? - The curious case of the Croatian football league

Tomislav Globan[a], Ed Jägers[b]

[a] *Faculty of Economics and Business, University of Zagreb*
[b] *DAT.Mobility, Deventer, the Netherlands*

November 20, 2016

Abstract

Participation of Lokomotiva Zagreb in the same tier of competition as Dinamo Zagreb has caused controversy ever since the former club entered the top tier of Croatian football league (1.HNL) in 2009. Claims that the two clubs are too closely connected in their sporting and business ventures raised suspicion in the fairness and regularity of the competition. This paper aims to determine if there are any matchups in 1.HNL that are characterized by the consistent over- or underperformance by one club against the other over the long run. If so, the study aims to test how significant these deviations are. To answer this question, we analyze over 5000 matches and all matchups that had ever taken place in 1.HNL, from seasons 1992 to 2015/16. We compare the realized number of points won by the teams in each matchup with the number of points each team was expected to win against the same opponent. To construct the expected number of points, we use the ClubElo formula that utilizes the Elo ranking system to measure the relative strength of teams. The results of our analysis indicate that one matchup extremely deviates from the expected distribution of points between the teams, and that matchup is Lokomotiva - Dinamo.

Keywords: Croatian football league, Elo ranking, ClubElo formula, football, sports economics, underperformance
JEL classification: L83, Z92

Authors wish to thank Lars Schiefler, the founder of clubelo.com, for his assistance with the data.
Authors contributed to this article in their personal capacity. The views expressed are their own and do not necessarily represent the views of the institutions they work for.

Contact address: Trg J. F. Kennedyja 6, 10000 Zagreb, Croatia. E-mail: tgloban@efzg.hr (Tomislav Globan); E-mail: ed.jagers@gmail.com (Ed Jägers).

1 Introduction

This paper aims to determine if there are any matchups in the Croatian top tier football league (1.HNL) that are characterized by the consistent long run over- or underperformance by one club against the other. If so, the study aims to test how significant these deviations are. Results of this analysis could prove very interesting with respect to currently one of the biggest controversies in Croatian football.

Ever since Lokomotiva Zagreb entered 1.HNL in 2009, a cloud of controversy has surrounded the organization. Their participation in the same level of competition as Dinamo Zagreb, the most successful club in Croatia and the club they have allegedly been closely connected to, prompted reactions from many stakeholders in Croatian football. For instance, Hajduk Split – historically the biggest domestic rival of Dinamo Zagreb – issued a formal request towards the Croatian football federation (CFF) to abstain from issuing Lokomotiva the license for participating in 1.HNL (Hajduk Split, 2015).

In their request, Hajduk claimed that Dinamo and Lokomotiva are too closely connected in their sporting and business ventures, and that Lokomotiva are directly dependent on Dinamo. They stated that as much as 50-70 percent of players that played for Lokomotiva over the previous five years had once been registered as Dinamo players and that the number of players that had previously been, or at the time were, on loan from Dinamo to Lokomotiva is unprecedentedly high (Hajduk Split, 2015). Other research has shown that no other club relationship in the league is characterized by even nearly the amount of player transfers and loans as the one between Dinamo and Lokomotiva (Tribina.hr, 2015).

This, according to Hajduk Split (2015), makes competition in 1.HNL unfair, and casts a shadow over the regularity of the competition. Civil association "Naš Hajduk" even issued a booklet with media excerpts, printscreens and document scans that, according to them, prove that Lokomotiva and Dinamo are two closely connected entities that form a "cartel", with Lokomotiva sometimes even being presented as Dinamo's "B team" (Naš Hajduk, 2014).

In 2014 the media reported that the newly formed Sports Inspection of the Ministry of Science, Education and Sports of the Republic of Croatia

issued a report in which they ordered CFF to retroactively revoke Lokomotiva the license for the season 2009/10 due to reasons "mainly related to the close sporting and business relationship between football clubs Dinamo and Lokomotiva" (Novi list, 2014). CFF immediately dismissed all allegations, disputed the report and maintained that the relationship between Lokomotiva and Dinamo is in accordance with the law and competition regulations which have been approved by UEFA (Hrvatski nogometni savez, 2014).

In this paper we do not analyze the closeness of the business relationship between Lokomotiva and Dinamo, nor aim to judge whether their simultaneous participation in 1.HNL has had any impact on the regularity and/or fairness of the competition. The sole purpose of the paper is to analyze the historical results of all matchups in the Croatian football league and see whether the results of the Lokomotiva - Dinamo matchup deviate in any way from other matchups.

To do so, we obtained data for over 5000 matches and over 850 matchups that had taken place in 1.HNL throughout all 25 seasons, from 1992 to 2015/16. We singled out 206 matchups suitable for the analysis and applied a three-step methodological approach. First, we calculate the sum of points each team was expected to win in each given matchup applying the Elo ranking principles and the ClubElo formula by Schiefler (2016). The formula provides historical *ex-ante* probabilities for each result (home-team win, a draw, away-team win) in each matchup, accounting for the relative strengths of teams, home advantage, and other factors. Then we relate expected points with points that have actually been realized by teams and calculate the historical realized-to-expected point ratios (REPR) to determine under- and overperformers and find outlier matchups. Finally, we test for the statistical significance of obtained values.

The results of our analysis indicate that one matchup extremely deviates from the expected distribution of points between the teams, and that is the Lokomotiva - Dinamo matchup. Lokomotiva have managed to collect only one point in 21 home and away matches against Dinamo in 1.HNL, which falls short of the expected performance by as much as 92%. No other matchup in the league comes even close to this level of underperformance. Its corresponding REPR value even lies outside the 99% confidence interval, far in the tail of normal distribution of REPR values.

To our knowledge, this is the first study of this kind. We have not found any similar matchup-by-matchup analysis of long run under- and overperformances within a league in any sport in the existing scientific literature.[1] This paper breaks new ground in that regard.

The rest of the paper is structured as follows. Section 2 explains the three-step methodology approach applied in the analysis and states the data sources. Results of the analysis, i.e. REPR values across matchups and significance tests, together with reports on robustness tests, are carried out in Section 3, while Section 4 concludes the paper.

2 Methodology and data

The methodological approach can be divided into three steps. First, the objective is to calculate the number of points each team was *ex-ante* expected to win in each given matchup. In the second step we calculate the historical realized-to-expected point ratios to determine under- and overperformers and find outlier matchups. Finally, we test for the statistical significance of obtained values.

2.1 Calculating the expected number of points

To calculate the number of points each team is expected to win in each matchup, we utilize the data from Schiefler (2016), i.e. the ClubElo website[2], which reports the historical *ex-ante* probabilities for each result (a home-team win, a draw, and an away-team win) in each given matchup for all European football leagues and continental competitions. The data for some leagues goes back to as far as 1939.

The ClubElo formula is based on the widely used Elo ranking system to measure the relative strengths of clubs in given points in time. Teams are ranked based on their Elo points, which generally increase with wins, and decrease with defeats. However, the formula takes into consideration the time component (with every new match played, the weights of previous matches decrease) and the strength of the opposition (win against a team with a high number of Elo points is valued more than a win against a team

[1] Perhaps worth mentioning is the study by Noland and Stahler (2016) who analyzed the (under)performance of Asian countries at the Olympic Games. However, that paper has very little common ground to our analysis here.

[2] Visit http://clubelo.com/ [accessed 7 November 2016].

with a low number of Elo points). The ranking system, originally used for the ranking of chess players, is modified to fit the club football purposes in that it also takes into account the goal difference in a match, i.e. a win by a higher margin brings more points than a win by a narrower margin. Table 1 reports the Elo rankings and points of teams playing in 1.HNL at the end of the 2015/16 season.

Table 1: Elo rankings of 1.HNL teams at the end of the 2015/16 season

Country rank	European rank	Club	Elo points
1	100	Dinamo Zagreb	1611
2	175	Rijeka	1530
3	345	Hajduk Split	1401
4	427	Lokomotiva Zagreb	1334
5	441	Slaven Belupo	1319
6	445	RNK Split	1316
7	459	Inter-Zaprešić	1307
8	491	Osijek	1261
9	536	NK Zagreb	1190
10	544	Istra 1961	1180

Source: clubelo.com

The probability distributions for a club's win, draw or defeat against the opposing team directly depend on the difference in Elo points between the two teams on the day the match is played. This means that the higher the Elo point difference (i.e. the difference in strength) between the teams, the higher the win probability of a better-ranked team is, and vice versa. Moreover, the home advantage is taken into account, as an algorithm[3] is used to adjust the Elo point difference slightly in favor of the home team. This increases the probability for a home-team win, as history has shown that teams are more successful when playing at home, rather than away.

Each Elo point difference (adjusted for the home advantage) has its own corresponding probability of a home-team win, a draw, and an away-team win. The match probabilities are calculated based on a result histogram for the corresponding Elo point difference. For instance, if team X has a 400 point home-adjusted advantage over team Y, the probability of a win by team X is determined by the historical percentage of wins by all other teams

[3] For details see http://clubelo.com/system

who had had the same Elo point advantage over their opponents. Probabilities for a draw or a win by team Y are obtained using the same principle.

Due to space limitations, it is not possible to report all details on this rather complex methodology of football club rankings and result probabilities, but one can easily find them on the ClubElo website (Schiefler, 2016).

2.2 Calculating the expected-to-realized point ratio

To determine the under- and overperforming matchups and single out extreme outliers we calculate the historical realized-to-expected point ratio (REPR) for all matchups ever played over 25 seasons of 1.HNL. Realized-to-expected point ratio of team *i* against team *j* is defined as:

$$REPR_{i,j} = \frac{realized_{i,j}}{expected_{i,j}} \tag{1}$$

where $realized_{i,j}$ denotes the realized total number of points won by team *i* against team *j*, both home and away, in all their matches throughout 25 seasons. Specifically, it is defined as:

$$realized_{i,j} = numW_{i,j} * 3 + numD_{i,j} * 1 \tag{2}$$

where $numW_{i,j}$ is the number of wins of team *i* over team *j*, while $numD_{i,j}$ is the number of draws between teams *i* and *j*. In football, a win is worth three points; a draw brings one point to both teams, while a defeat is worth zero points.

On the other hand, $expected_{i,j}$ denotes the total number of points team *i* was expected to win against team *j*, both home and away, in all their matches throughout 25 seasons. Specifically, the following formula was used:

$$expected_{i,j} = \sum_{n=1}^{p} \frac{(prW_n * 3 + prD_n * 1)}{100} \tag{3}$$

where prW_n is an *ex-ante* percent probability of team *i* winning over team *j* on the day the match *n* is played; prD_n is an *ex-ante* percent probability of a draw between teams *i* and *j* on the day the match *n* is played, while *p* is the total number of matches played between teams *i* and *j* over 25 seasons of 1.HNL.

All matches ever played in 1.HNL, from its inception in spring 1992[4] to the end of season 2015/16, have been considered for the analysis. The data source for these over 5000 match results was the official website[5] of 1.HNL (MAXtv Prva liga, 2016). The total of 36 clubs have taken part in 1.HNL over this period of time. Theoretically, this provides the potential total of 1260 matchups.[6] However, not all potential matchups have actually taken place, because the structure of the league changes with each season, as bottom teams are relegated to the lower tier competition (2.HNL), and the best ones from 2.HNL are promoted to 1.HNL. Thus, in reality, 866 matchups have taken place over the analyzed period of time.

Furthermore, not all matchups that have actually taken place are considered for the calculations of REPR. This is because a low number of matches between two clubs could create biased REPR values. For instance, let us consider two cases. In the first one, teams X and Y have played only two matches throughout the analyzed period of time, and both were won by team Y. Here, the value of $REPR_{X,Y}$ equals zero. In the second case, teams X and Y have played as many as 20 matches throughout the analyzed period of time, with all of them also ending with a win by team Y. Here, the value of $REPR_{X,Y}$ equals zero, just like in the first case. However, the two values are not directly comparable, because the second case represents a much bigger underachievement by team X than the first one. The probability of winning zero points in 20 matches against the same opponent is extremely low, whereas the same point tally in only two matches is much more probable. For this reason, the cut-off minimum number of matches between the clubs was set to 15. This decreased the number of matchups entering the analysis to 206. Refer to Table 3 in the Appendix for the total number of matches per each matchup and to see which matchups entered the analysis.

It should be noted that the matchups "team X vs. team Y" and "team Y vs. team X" are considered as separate matchups, although they cover the same set of home and away matches between the clubs. This is because the $REPR_{i,j}$ values (realized and expected points) differ depending on which club is considered as team i and which one as team j.

[4] Due to the war activities in Croatia in autumn 1991, the whole 1991/92 season was played in the spring of 1992. The usual schedule involves one half-season played in autumn, and the other in spring of the following year.

[5] Visit http://prvahnl.hr/povijest/rezultati-i-poretci/ [accessed 7 November 2016].

[6] The formula is $k(k-1)$, where k is the number of participating clubs.

2.3 Testing the significance of REPR values

A z-score is assigned to every matchup based on its REPR value, which is then used to calculate the p-value per matchup, i.e. to test the probability that the observed REPR is significantly different from the sample average (μ). This way it is possible to find out if there are any opponents that each team significantly under- or overperforms against.

In essence, we test whether the null hypothesis of $REPR_{i,j} = \mu$ can be rejected at the usual levels of significance. In case of statistically significant values of $REPR_{i,j} < \mu$, team i can be considered as an underperformer against team j, as it won less points than expected based on the relative strengths of teams. On the other hand, in case of statistically significant values of $REPR_{i,j} > \mu$, team i can be considered as an overperformer against team j, as it won more points than expected.

3 Results

First, we deal with the question who are, if any, the opponents each team under- and overperforms against, respectively. To that end, we report the calculated realized-to-expected point ratios per each matchup (Table 2), together with their respective significance test results. REPR values were calculated based on realized and expected number of points for each matchup, which are reported in Table 4 in the Appendix.

Red-shaded cells in Table 2 indicate matchups in which team i significantly underperforms against team j, while green-shaded cells denote that team i significantly overperforms against team j. Cells left unshaded represent matchups in which team i performed according to expectations against team j, as the matchup's REPR value does not significantly deviate from the sample mean at usual significance levels (1% and 5%).

Figure 1 sums the performance distribution of matchups per each club. It is visible that the relative number of underperforming matchups is lowest for Dinamo Zagreb. Only 7% of all their matchups are classified as an underperformance, and that is the one against Hajduk Split, their historically biggest archrivals. The next in line is Lokomotiva Zagreb with only 11% of matchups considered as an underperformance (the one against Dinamo Zagreb). On the other side of the spectrum, the rate of underperformances is the highest for Kamen Ingrad and Istra 1961, with 63% and 62%, respectively.

Figure 1: Distribution of total matchups per club by performances

The highest relative number of overperformance matchups can be found at Lokomotiva (67%), followed by Rijeka (63%). Cibalia and Hrvatski dragovoljac have the lowest relative number of overperformances in the league (29% and 30%, respectively).

Next, we aim to differentiate between the levels of under- and overperformance and identify the biggest outliers amongst these matchups. To that end, we have calculated the respective lower and upper boundaries for non-extreme REPR values in the sample. Specifically, we utilize the fact that 95% of observed values in a standard normal distribution lie within the $\langle \mu - 1.96\sigma, \mu + 1.96\sigma \rangle$ range, where μ is a sample mean and σ a sample standard deviation.

Given that in our sample $\mu = 0.990$ and $\sigma = 0.173$, the lower 95% confidence band is set at the REPR value of 0.650, and the upper band at 1.329. All matchups with REPR values below 0.650 or above 1.329 are outliers, i.e. could be viewed as matchups where team *i* extremely underperforms (or in the latter case extremely overperforms) against team *j*.

Table 2: Calculated REPR values

team *i* \ team *j*	Cibalia	Dinamo Zagreb	Hajduk Split	Hrv. dragovoljac	Inter-Zaprešić	Istra 1961	Kamen Ingrad	Lokomotiva	Marsonia	NK Zagreb	Osijek	Rijeka	RNK Split	Šibenik	Slaven Belupo	Varteks	Zadar
Cibalia		0.630[a]	0.981	0.996	1.045[a]	0.962[b]	0.850[a]		0.681[a]	0.995	0.986	1.033[a]		1.147[a]	1.002	1.077[a]	0.733[a]
Dinamo Z.	1.109[a]		0.947[a]	1.051[a]	0.971	1.030[a]	1.110[a]	1.325[a]		0.990	1.121[a]	1.010	1.054[a]	1.025[a]	1.098[a]	0.983	0.971
Hajduk S.	0.960[b]	1.090[a]		0.852[a]	1.041[a]	0.885[a]	0.941[a]	1.005		1.060[a]	1.082[a]	0.950[a]	1.091[a]	1.164[a]	1.102[a]	1.137[a]	1.093[a]
Hrv. drag.	0.964[b]	0.804[a]	1.357[a]							0.872[a]	1.245[a]	0.634[a]		0.974	0.786[a]	0.709[a]	1.143[a]
Inter-Z.	0.926[a]	1.106[a]	0.843[a]			1.175[a]				0.893[a]	0.620[a]	1.249[a]		0.870[a]	1.179[a]	0.951[a]	1.087[a]
Istra 1961	1.062[a]	0.860[a]	1.326[a]		0.866[a]			0.920[a]		0.736[a]	0.913[a]	0.893[a]	1.170[a]	0.987	1.018[b]	0.962[b]	0.913[a]
Kamen In.	1.109[a]	0.594[a]	1.315[a]							0.928[a]	0.725[a]	0.935[a]			1.121[a]	0.956[a]	
Lokomotiva		0.079[a]	1.052[a]			1.159[a]				1.420[a]	1.303[a]	0.976	0.974		1.043[a]		1.153[a]
Marsonia	1.332[a]											0.923[a]					
NK Zagreb	0.952[a]	0.903[a]	1.040[a]	0.999	1.066[a]	1.199[a]	1.079[a]	0.638[a]			0.879[a]	1.009	0.779[a]	1.024[a]	0.765[a]	0.979	1.120[a]
Osijek	1.032[a]	0.608[a]	0.898[a]	0.825[a]	1.307[a]	1.026[a]	1.400[a]	0.600[a]		1.115[a]		1.079[a]	0.874[a]	1.070[a]	0.801[a]	1.142[a]	1.014[b]
Rijeka	0.951[a]	1.005	1.076[a]	1.297[a]	0.819[a]	1.052[a]	1.201[a]	1.022[a]	1.186[a]	1.035[a]	0.928[a]		1.049[a]	1.051[a]	0.962[b]	1.050[a]	0.958[a]
RNK Split		0.934[a]	0.693[a]			0.830[a]		0.886[a]		1.286[a]	1.041[a]	0.932[a]			1.264[a]		1.068[a]
Šibenik	0.874[a]	0.902[a]	0.633[a]	1.049[a]	1.159[a]	0.998				1.036[a]	0.886[a]	0.937[a]			0.808[a]	0.952[a]	1.114[a]
Slaven B.	0.963[b]	0.733[a]	0.722[a]	1.017[b]	0.899[a]	0.929[a]	0.958[a]	0.900[a]		1.281[a]	1.218[a]	0.935[a]	0.713[a]	1.139[a]		1.110[a]	1.101[a]
Varteks	1.061[a]	1.126[a]	0.801[a]	1.158[a]	1.103[a]	1.045[a]	1.109[a]			1.017[b]	0.885[a]	0.824[a]		1.010	0.900[a]		1.045[a]
Zadar	1.297[a]	1.034[a]	0.779[a]	0.845[a]	1.011	1.101[a]		0.906[a]		0.843[a]	1.027[a]	1.049[a]	0.952[a]	0.790[a]	0.980	0.963[b]	

Note: Significant at: [a] 1 percent, [b] 5 percent. Red-shaded cells denote that team *i* significantly underperforms against team *j*. Green-shaded cells denote that team *i* significantly overperforms against team *j*. Some cells are empty because less than 15 matches were played between the clubs.

Judging by the histogram (Figure 2), calculated REPR values seem normally distributed, apart from the one matchup to the far left, with a REPR value lower than 0.1. This is the matchup Lokomotiva - Dinamo, with the REPR value of 0.079.

Figure 2: Histogram of calculated REPR values

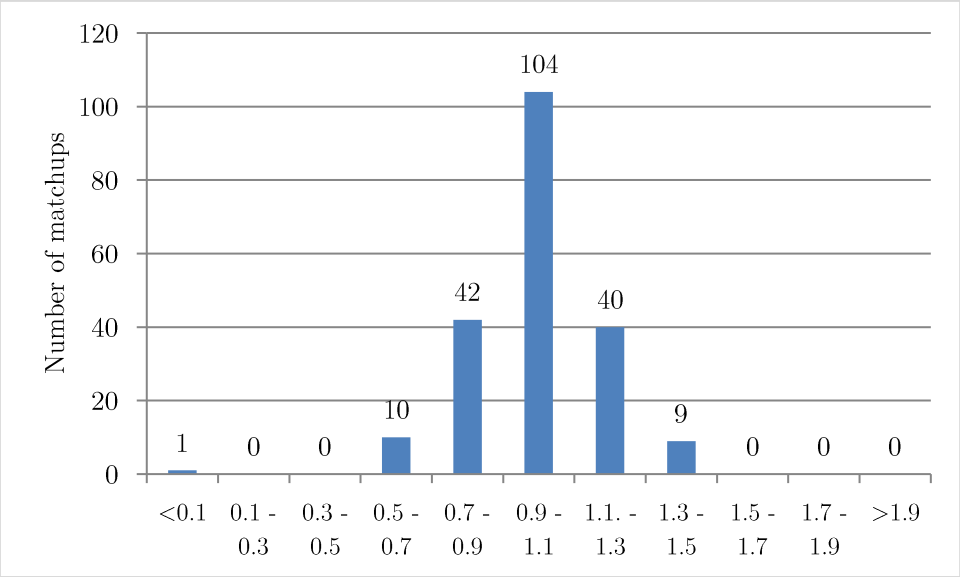


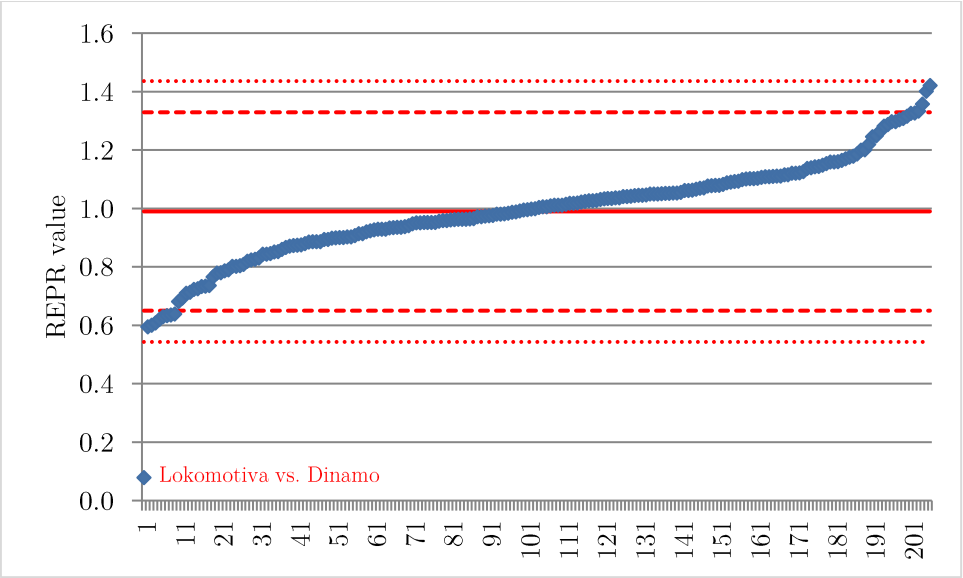
Figure 3 orders all matchups by ascending REPR values and alphabetically by club names, along with 95% and 99% confidence bands[7]. Out of the total 206 matchups, 193 of them lie within the 95% confidence interval. This leaves 13 matchups that can be considered as outliers, with nine of them categorized as extreme underperformances and four as extreme overperformances.

Only one matchup is well beyond even the 99% confidence interval, and that is, again, Lokomotiva - Dinamo. In 21 matches against Dinamo in 1.HNL, home and away, Lokomotiva managed to collect only one point (0 wins, 1 draw, 20 defeats), while the expected point tally was 12.79, which is the performance of 92% below expectation considering the relative strength of clubs. No other matchup comes even close to this level of underperformance, as is evident from Figure 3.

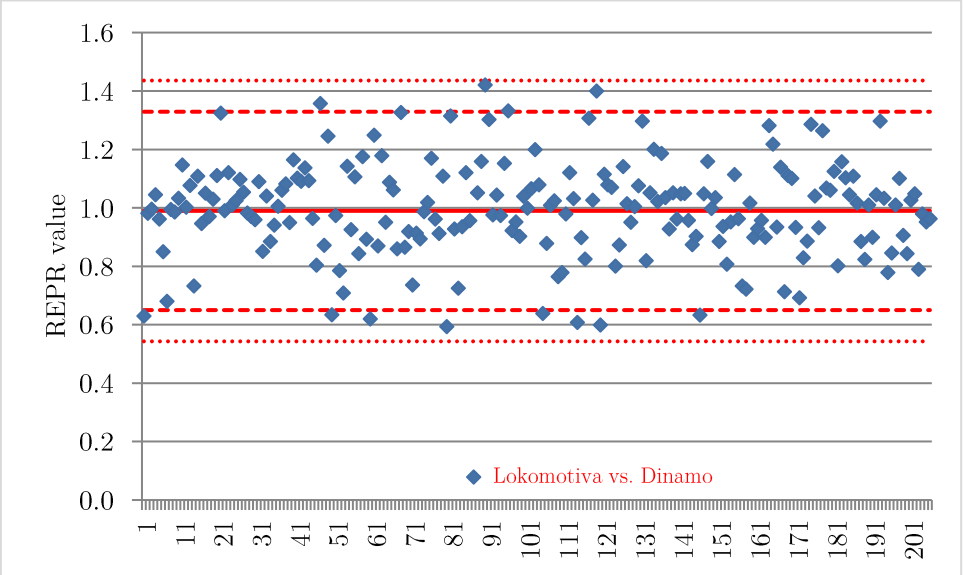
[7] The lower 99% confidence band is set at the REPR value of 0.543, and the upper band at 1.436.

Figure 3: All 206 matchups ordered

(a) by ascending REPR values



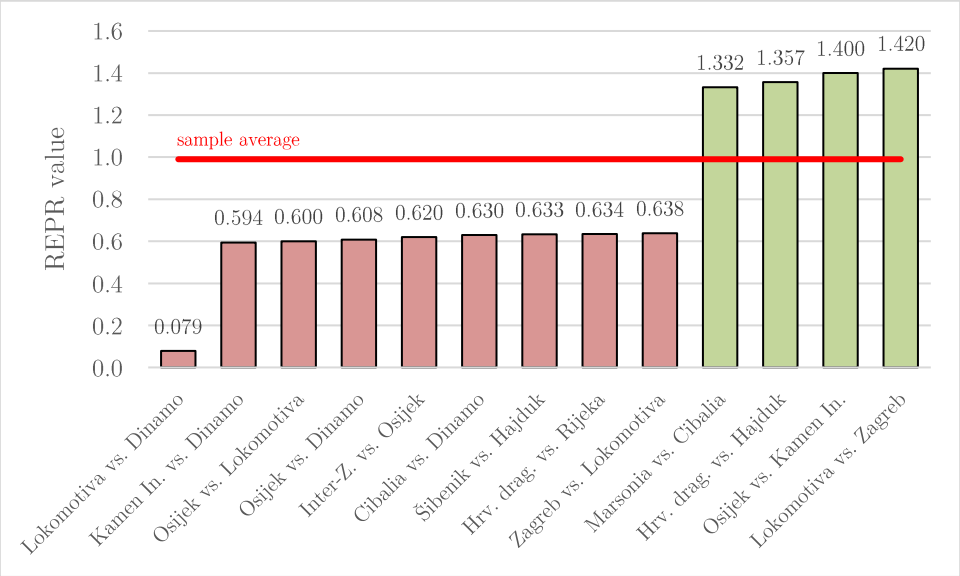
(b) alphabetically by club names



Note: Full red line represents the sample average, dashed lines represent 95% confidence bands, dotted lines represent 99% confidence bands.

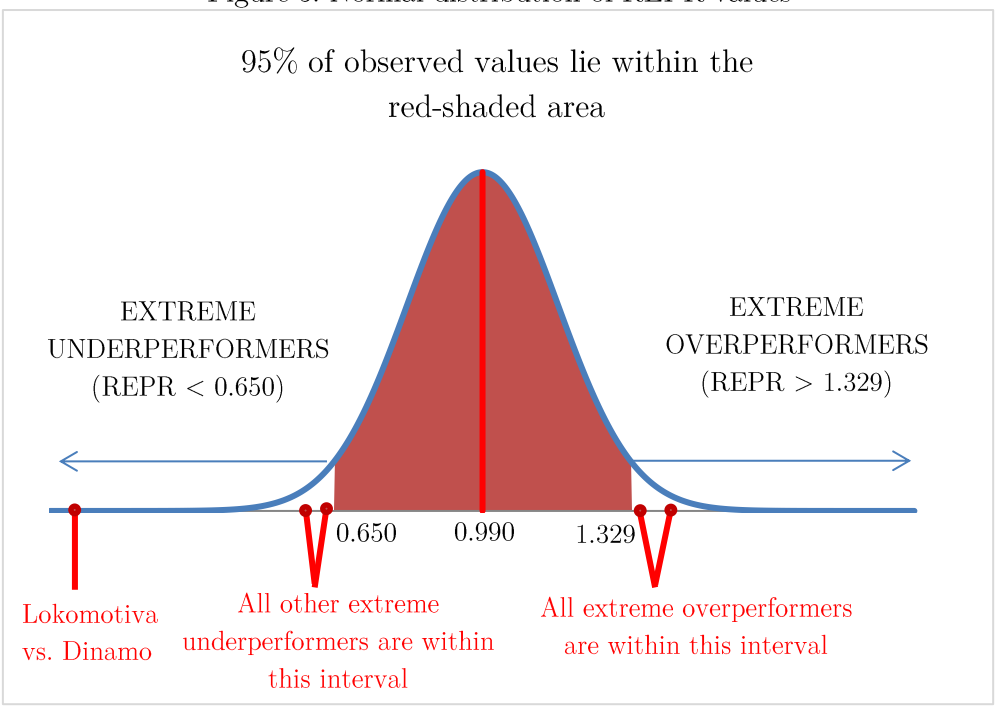
Figure 4 reveals the details on other outlier matchups. It is interesting to note that seven out of nine extreme underperformance matchups involve at least one of Dinamo, Lokomotiva or Osijek.

Figure 4: Extreme under- and overperformance matchups



Finally, Figure 5 shows where most matchups lie within the normal distribution; where all outliers, except for one, are located, and confirms how far in the tail of the distribution (i.e. how improbable) the Lokomotiva - Dinamo matchup actually is.

Figure 5: Normal distribution of REPR values



These findings are even more unusual given the fact that Lokomotiva have not underperformed against any other club in the league against whom they played at least 15 matches (see Table 2). Quite contrary, they have recorded a statistically significant overperformance in 67% of their total matchups (Figure 1), a feat matched by no other club in the league history. Furthermore, Lokomotiva have continuously proven to be a member of the upper quality class of Croatian football. In their only seven seasons in the top tier of Croatian football league, they were once vice-champions (2012/13), placed four times in the top 5, and for six out of seven seasons they finished in the upper half of the table. They qualified for European competitions several times and Elo ranking placed them as the fourth strongest Croatian club at the end of the 2015/16 season (Table 1). All this makes the aforementioned findings even more surprising.

To test the robustness of obtained results, we made two changes in our methodology. First, we calculated REPR values for the case where teams are awarded two instead of three points for a win. One could argue that the "three point rule"[8] may artificially inflate or deflate REPR values, because winning and losing is worth more than drawing twice, while in fact both scenarios mean that teams are on average equally good. To test whether this had any impact on our results, we changed equations (2) and (3) in that $numW_{i,j}$ and prW_n were respectively multiplied by 2. Although the REPR value for the matchup Lokomotiva - Dinamo increased from 0.079 to 0.101, it remained by far the biggest outlier amongst all 206 matchups, and the only one significant at the 1% level. Furthermore, all of the findings presented previously remained intact as a result of this change.

The second change deals with the cut-off number of matches between the clubs for the matchup to be included in REPR calculations. One could argue that the threshold set at 15 matches is still too low to avoid biased values of REPR. Thus, we increased the cut-off number to the highest possible that includes the matchup Lokomotiva - Dinamo, and that is 21. This decreased the total number of matchups entering the analysis to 152. Again, this re-

[8] This rule was officially introduced by FIFA in 1994 for the World Cup, after which all other competitions followed the practice. The rationale for the change was the aim to encourage more attacking play, because teams would be less willing to settle for a draw if they had the possibility to gain two extra points by scoring a winning goal, which outweighs the prospect of losing one point by conceding a losing goal.

sulted in no changes in the aforementioned findings. Detailed results of robustness checks are not reported due to space issues, but are available upon request.

4 Conclusion

The purpose of this paper was to determine if there are any clubs in the Croatian top tier football league that consistently over- or underperform against same opponents and how significant these deviations are. To our knowledge this is the first research of this kind. To do so, we obtained data for over 5000 matches and all matchups that have ever taken place in 1.HNL throughout 25 seasons since the competition's inception in 1992. We singled out 206 matchups suitable for the analysis and applied a three-step methodological approach.

In the first step we calculated the number of points each team was *ex-ante* expected to win in each given matchup. To construct the expected number of points for each team in each matchup, we used the ClubElo formula that utilizes the widely used Elo ranking system to measure the relative strength of teams. In the second step we analyzed and compared the realized number of points won by teams in each matchup with the number of points each team was expected to win against the same opponent. To do so, we calculated the historical realized-to-expected point ratios to determine under- and overperformers and find outlier matchups. Finally, we tested for the statistical significance of obtained REPR values.

The results of our analysis indicate that one matchup significantly deviates from the expected distribution of points between the teams, and that is the Lokomotiva - Dinamo matchup. Lokomotiva have managed to collect only one point in 21 home and away matches against Dinamo in 1.HNL, which falls short of the expected performance by 92%. No other matchup in the league comes even close to this level of underperformance. Its corresponding REPR value even lies outside the 99% confidence interval, far in the tail of normal distribution of REPR values. Our results are robust across various specifications of the model.

These findings seem even more unusual given the fact that Lokomotiva have not underperformed against any other club in the league, have the

highest overperformance rate in the competition's history, and have continuously proven to be a member of the upper quality class of Croatian football by their rankings in the competition.

Lokomotiva's participation in 1.HNL has been surrounded by a cloud of controversy ever since it entered the top tier of the Croatian football league. The reason are the claims that they are too closely connected to Dinamo Zagreb in their sporting and business ventures (even dubbed as "a cartel" by some), raising suspicion in the fairness and regularity of the competition where both teams play against each other.

Although the findings of this paper certainly will not decrease the levels of suspicion regarding the issue, this research in no capacity aims to answer the question whether the simultaneous participation of Lokomotiva and Dinamo in 1.HNL makes the competition irregular or unfair, nor to suggest that any actions should or should not be taken in this matter.

References

Hajduk Split (2015). *Evo zasto smo trazili da se Lokomotivi ne izda licenca.* [online] Available at: http://hajduk.hr/vijest/evo-zasto-smo-trazili-da-se-lokomotivi-ne-izda-licenca/5825 [Accessed 14 July 2016].

Hrvatski nogometni savez (2014). *Odgovor HNS-a na rješenje sportske inspekcije*. [online] Available at: http://hns-cff.hr/news/4391/odgovor-hns-a-na-rjesenje-sportske-inspekcije/ [Accessed 14 November 2016].

MAXtv Prva liga (2016). [online] Available at: http://prvahnl.hr/ [Accessed 12 November 2016].

Naš Hajduk (2014). *Kartel*. 1st ed. [pdf] Available at: http://hajduk.hr/pdf/nas-hajduk-kartel/25 [Accessed 14 July 2016].

Noland, M., and Stahler, K. (2016). Asian participation and performance at the Olympic Games. *Asian Economic Policy Review*, 11(1), 70-90.

Novi list (2014). *HNS u roku od 15 dana mora izbaciti Lokomotivu iz lige!*. [online] Available at: http://www.novilist.hr/Sport/Nogomet/HNS-u-roku-od-15-dana-mora-izbaciti-Lokomotivu-iz-lige?meta_refresh=true [Accessed 14 November 2016].

Schiefler, L. (2016). Football Club Elo Ratings. [online] Available at: http://clubelo.com/ [Accessed 12 November 2016].

Tribina.hr (2015). *Tribina.hr istraživanje: Lokomotiva je Dinamova B-momčad? DA*. [online] Available at: https://www.tribina.hr/topics/slug/

tribinahr-istrazivanje-lokomotiva-je-dinamova-b-mo/ [Accessed 14 November 2016].

Appendix

Table 3: Total number of matches (home and away) between clubs that participated in 1.HNL between 1992 and 2015/16

	Belišće	Čakovec	Cibalia	Croatia S.	Dinamo Z.	Dubrava	Dubrovnik	Hajduk S.	Hrv. drag.	Inter-Z.	Istra 1961	Kamen In.	Karlovac	Lokomotiva	Lučko	Marsonia	Međimurje	Mladost 127	Neretva	NK Zagreb	Orijent	Osijek	Pazinka	Pomorac	Primorac	Radnik	Rijeka	RNK Split	Samobor	Segesta	Šibenik	Slaven B.	TŠK Top.	Varteks	Vukovar 91	Zadar
Belišće			6		6	2	8	6	2	6	6					2		2	6	6	4	6	4		8	4	6			6	6			6		10
Čakovec			6		4			4	6			2				6				4		4		2			6				6	4	2	4		2
Cibalia	6	6		5	47	2	6	47	15	37	33	15	6	9	2	18	12	8	2	53	2	45	4	4	4	4	49	7		12	44	39	2	44	3	35
Croatia S.			5		5			5		5	2		2	2			2			5		5					5				5	5		5		5
Dinamo Z.	6	4	47	5		2	6	74	25	37	45	19	6	21	2	12	14	6	2	65	2	75	4	4	4	4	67	19	2	10	40	51	2	58	3	51
Dubrava	2		2		2		2	2		2	2									2		2	2		2	2	2			2	2			2		2
Dubrovnik	8		6		6	2		6	2	6	6							2	4	6	4	6	4		6	4	6			4	6			6		10
Hajduk S.	6	4	47	5	74	2	6		25	39	43	17	6	21	2	12	12	6	2	67	2	73	4	4	4	4	69	19	2	10	40	53	2	60	3	49
Hrv. drag.	2	6	15		25		2	25		4	11	2	2	6		8		8	2	19	4	25		2	2		23	6	2	2	17	17	2	19	3	18
Inter-Z.	6		37	5	37	2	6	39	4		33	8	6	13	2	12	11	4	2	43	2	37	4		4	4	41	11		12	26	31		32		29
Istra 1961	6		33	2	45	2	6	43	11	33		9	6	21	2	8	13	4	2	41	2	45	4		4	4	45	19		12	26	33		28	3	31
Kamen In.		2	15		19			17	2	8	9					6	9			17		17		4			15				7	15	2	17		10
Karlovac			6	2	6			6	2	6	6			6	2		2			6		6					6	4			6	6		6		6
Lokomotiva			9	2	21			21	6	13	21		6		2		2			17		21					21	19			6	21		6		17
Lučko			2		2			2		2	2		2	2						2		2					2	2			2	2		2		2
Marsonia	2	6	18		12			12	8	12	8	6						4	2	14	2	12		2	2		16			8	14	8	2	12		8
Međimurje			12	2	14			12		11	13	9	2	2						14		14					12				8	14		12		9
Mladost127	2		8		6		2	6	8	4	4					4			2	8	4	6			2		10		4	4	12	8		8		10
Neretva	6		2		2		4	2	2	2	2					2		2		2	4	2			6		2			2	2			2		6
NK Zagreb	6	4	53	5	65	2	6	67	19	43	41	17	6	17	2	14	14	8	2		2	63	4	4	4	4	59	15	2	10	42	55	2	54	3	45
Orijent	4		2		2		4	2	4	2	2					2		4	4	2		2			4		2			2	2			2		6
Osijek	6	4	45	5	75	2	6	73	25	37	45	17	6	21	2	12	14	6	2	63	2		4	6	4	4	69	19	2	10	42	53	2	56	3	53
Pazinka	4		4		4	2	4	4		4	4									4		4			2	4	4			4	4			4		4
Pomorac		2	4		4			4	2			4				2				4		6					6				6	6	2	4		6
Primorac	8		4		4	2	6	4	2	4	4					2		2	6	4	4	4	2			2	4			4	4			4		8
Radnik	4		4		4	2	4	4		4	4									4		4	4		2		4			4	4			4		4
Rijeka	6	6	49	5	67	2	6	69	23	41	45	15	6	21	2	16	12	10	2	59	2	69	4	6	4	4		19	4	12	48	55	2	56	3	49
RNK Split			7		19			19	6	11	19		4	19	2					15		19					19				4	19		4		15
Samobor					2			2	2									4		2		2					4				4	4		4		2
Segesta	6		12		10	2	4	10	2	12	12					8		4	2	10	2	10	4		4	4	12				12			10		8
Šibenik	6	6	44	5	40	2	6	40	17	26	26	7	6	6	2	14	8	12	2	42	2	42	4	6	4	4	48	4	4	12		34	2	42	3	34
Slaven B.		4	39	5	51			53	17	31	33	15	6	21	2	8	14	8		55		53		6			55	19	4		34		2	40	3	39
TŠK Top.		2	2		2			2	2			2				2				2		2		2			2				2	2		2		2
Varteks	6	4	44	5	58	2	6	60	19	32	28	17	6	6	2	12	12	8	2	54	2	56	4	4	4	4	56	4	4	10	42	40	2		3	36
Vukovar 91			3		3			3	3		3									3		3					3				3	3		3		
Zadar	10	2	35	5	51	2	10	49	18	29	31	10	6	17	2	8	9	10	6	45	6	53	4	6	8	4	49	15	2	8	34	39	2	36		

Note: Matchups entering the analysis are shaded gray.

Table 4: Realized and expected points per matchup

		Cibalia	Dinamo Z.	Hajduk S.	Hrv. drag.	Inter-Z.	Istra 1961	Kamen In.	Lokomotiva	Marsonia	NK Zagreb	Osijek	Rijeka	RNK Split	Šibenik	Slaven B.	Varteks	Zadar
Cibalia	realized		17	37	20	54	45	17		19	61	50	58		71	46	61	38
	expected		27.00	37.71	20.09	51.66	46.76	19.99		27.91	61.33	50.70	56.13		61.90	45.89	56.63	51.87
Dinamo Z.	realized	116		118	61	82	105	46	61		136	176	134	42	94	122	120	119
	expected	104.64		124.60	58.05	84.46	101.94	41.43	46.05		137.31	157.03	132.65	39.85	91.66	111.12	122.09	122.54
Hajduk S.	realized	88	85		42	80	75	32	37		130	141	105	33	94	107	127	111
	expected	91.66	78.01		49.32	76.84	84.73	34.02	36.81		122.59	130.29	110.50	30.24	80.76	97.10	111.73	101.54
Hrv. drag.	realized	20	10	27							20	34	16		22	18	16	29
	expected	20.74	12.44	19.90							22.93	27.31	25.22		22.59	22.91	22.57	25.38
Inter-Z.	realized	45	22	26			55				46	27	57		30	43	39	46
	expected	48.61	19.89	30.84			46.80				51.51	43.55	45.65		34.49	36.48	41.00	42.30
Istra 1961	realized	45	21	45		37			24		36	48	40	25	32	42	31	42
	expected	42.39	24.42	33.93		42.75			26.10		48.92	52.55	44.77	21.37	32.42	41.24	32.22	46.01
Kamen In.	realized	23	7	17							19	16	15			21	17	
	expected	20.74	11.79	12.93							20.47	22.08	16.04			18.74	17.78	
Lokomotiva	realized		1	22			36				35	40	20	24		31		32
	expected		12.72	20.91			31.05				24.64	30.70	20.50	24.64		29.71		27.76
Marsonia	realized	28											16					
	expected	21.02											17.34					
NK Zagreb	realized	79	40	64	29	70	75	28	14			74	72	13	68	55	75	85
	expected	83.00	44.32	61.54	29.03	65.67	62.54	25.95	21.94			84.20	71.38	16.69	66.40	71.91	76.58	75.88
Osijek	realized	74	32	63	34	75	72	34	16		98		91	19	72	59	95	89
	expected	71.69	52.66	70.16	41.21	57.39	70.15	24.28	26.68		87.91		84.38	21.75	67.29	73.66	83.22	87.78
Rijeka	realized	73	53	84	49	54	82	30	38	31	93	97		32	79	85	89	83
	expected	76.79	52.76	78.05	37.79	65.90	77.95	24.99	37.18	26.13	89.81	104.54		30.51	75.14	88.38	84.79	86.65
RNK Split	realized		12	15			25		24		31	31	20			35		27
	expected		12.85	21.66			30.13		27.09		24.10	29.79	21.45			27.69		25.27
Šibenik	realized	50	19	19	25	42	38				50	42	52			35	49	55
	expected	57.19	21.06	30.01	23.84	36.24	38.08				48.28	47.43	55.49			43.33	51.48	49.35
Slaven B.	realized	58	23	35	24	43	45	21	25		100	86	58	17	56		61	68
	expected	60.26	31.38	48.46	23.61	47.82	48.45	21.92	27.79		78.03	70.58	62.06	23.85	49.17		54.94	61.78
Varteks	realized	67	45	43	34	51	46	32			72	62	56		64	49		62
	expected	63.17	39.98	53.65	29.36	46.24	44.03	28.87			70.80	70.09	67.98		63.34	54.46		59.34
Zadar	realized	56	23	27	20	37	42		17		40	59	50	15	34	44	38	
	expected	43.18	22.25	34.68	23.66	36.62	38.14		18.76		47.45	57.46	47.67	15.76	43.05	44.91	39.44	