STOCKHOLM SCHOOL OF ECONOMICS
Department of Economics
659 Degree project in economics
Spring 2015

KEYS TO FOOTBALL SUCCESS

An economic study of external factors affecting the football performance of Swedish municipalities

Joakim Kjeller (22853)

Abstract: Despite the popularity of football, no one has made a study of what external factors affect the chances of being a successful football city in Sweden. In this thesis a statistical study, using a Fixed Effects model, of all Swedish municipalities in the years 2003-2012 is performed to identify the most important external factors affecting the football performance. The number of points in Allsvenskan, the highest Swedish football division, is used as the measure of performance and the explanatory variables are chosen based on previous research. The results show that the optimal Swedish football city has a recent history of being successful in football and it is unlikely that we will see many new municipalities represented in Allsvenskan, in the future. Increases in the educational level also have a small positive impact on the football performance. Competition from an additional team in Allsvenskan from the same county, but from another municipality, has a small negative impact on the football performance of a municipality.

Keywords: football, Allsvenskan, Swedish municipalities, Fixed Effects model

JEL: Z20

Supervisor: Kelly Ragan
Date submitted: 17 May 2015
Date examined: 5 June 2015
Discussant: Jesper Vinge
Examiner: Karl Wärneryd

Table of contents

1 Introduction

1.1 Background

Football is one of the biggest sports in Sweden. The latest published report from Swedish Sports Confederation shows that football is the sport with most members (Swedish Sports Confederation 2014, p. 64) and the year before the same organization published statistics showing that football was the sport with most spectators in Sweden (Swedish Sports Confederation 2013, p. 61).[1] The highest Swedish football division, Allsvenskan, is the national league with the highest average attendance and the second division, Superettan, is number four in the same list (Sveriges Television 2014). The clubs in Allsvenskan have had a yearly turnover of around SEK 1.25 billion the last years (Swedish Football Association 2014, p. 8) and it is easy to see that football has a big impact on society, both economically and in other aspects. Football affects society and society affects football. Despite that, no one has made a study of external factors affecting the chances of success on the football pitch, for Swedish cities.

1.2 Purpose

The purpose of this thesis is to determine what external factors are affecting how successful a Swedish city is on the football pitch. The results obtained should be possible to use to investigate if a city over- or underperforms in terms of football. From that, it should be possible to make predictions about which cities that may improve in terms of football, in the long run, and about which cities that can expect darker times to come. The most important purpose of this thesis is, however, to point out areas for further research. If, for example, it is shown that population has a big impact on a city's football performance, the next step is to investigate why. This research can be undertaken by other economists or by researchers within other scientific areas.

1.3 Problem specification

It is hard to find statistics for cities, as cities do not have any legal status since the municipal reform 1971 (Cederfelt 2012). Therefore municipalities have been evaluated in this thesis.[2] Localities could have been an alternative, but the localities are so many that the workload would have been too big.[3]

To limit the scope and to avoid a situation in which success in different divisions needs to be compared, the focus is only on Allsvenskan, so lower divisions will not be taken into account. International achievements and the international status of Allsvenskan will also be overlooked, since only domestic success is evaluated.

The years in focus of this thesis are 2003-2013 and there are mainly two reasons why not a longer period has been studied. First and foremost, data for all the variables has not been

[1] The report covers the year 2013.
[2] "Kommun" is the Swedish word for municipality.
[3] "Tätort" is the Swedish word for locality.

possible to find for more years. Especially the climate data is limiting, since data is only publicly available from 2002 and onwards. The political party Sverigedemokraterna is also limiting, as the party only has been reported separately in the statistics since the 2002 election. The second reason is that the football has changed quite dramatically, so years earlier than 2003 may not be representative for the situation today.[4] 2013 is used to evaluate the accuracy of the estimation constructed by the regression and therefore 2013 is not included in the regression.

Municipalities can be thought of as utility maximizing units. The clubs in a municipality use their external factors to produce as good output as possible, in this case successful football teams.[5] The output can be measured in a couple of ways[6], but my opinion is that the most accurate measure is number of points[7]. The output of every single game is points that are added to a table, so the aim of every single game is to generate points. Of course the points taken affect a team's position in the table, but the table position is also dependent on how other teams perform. For example, Helsingborgs IF took 65 points 2010, but ended second after Malmö FF with 67 points (Alsiö 2011, p. 506). They had a better season in terms of match results than IF Elfsborg had 2012 when they won Allsvenskan at 59 points (Swedish Football Association 2012). Number of points is the best measure to compare long run football performance. Furthermore, points are easier to add for a municipality than, for example, table positions.[8] This thesis will not evaluate how the output, in terms of points, is translated into utility, but it is reasonable to expect that one point in Allsvenskan translates into different amounts of utility for different municipalities.

The model used to establish the perceived utility for a municipality can be written as follows:
$U=U\{P[I_1, I_2, \ldots I_k, E_1, E_2, \ldots E_k]\}$
U is the perceived utility for a municipality and depends on P, the number of points in Allsvenskan for all the clubs from the municipality. P is in turn affected by the internal usage $I_1, I_2, \ldots I_k$ of the external factors $E_1, E_2, \ldots E_k$.[9] The aim of this thesis is to identify $E_1, E_2, \ldots E_k$ and to evaluate their effects on P, given average internal resource usage. The research question is:
What external factors affect a Swedish municipality's chances of being successful in football, and in which ways?

[4] For example, the football has changed in terms of commercialization.
[5] External factors are outside factors influencing the football performance, that cannot be controlled by the clubs from a municipality.
[6] For example, football performance can be measured in terms of titles, lower/upper half of the table, and so on.
[7] The winner of a football game gets three points and the loser gets zero points. If a game ends up in a draw, the teams get one point each.
[8] See "Points in Allsvenskan" under "Football performance in Sweden" in the "Variable creation" section in the "Appendix" for a description of how this variable has been formed.
[9] The internal usage takes place in a municipality's football clubs and affects the output of the municipality, in terms of points.

To be able to tackle the question, a hypothesis is formulated. The hypothesis is that seven key external factors, identified by analysing related research, have a statistically significant impact on the football performance of a Swedish municipality.[10]

2 Current state of knowledge and variable selection

2.1 Identification criteria

By doing a literature study of related research, it is possible to identify seven key external factors that may affect the football performance of a Swedish municipality. These areas are presented in the following subsections and form the basis of this entire thesis. Since the field of research in which this thesis takes its starting point is rather unexplored, some of the sources in the "Current state of knowledge and variable selection" section are referred to many times. This field of research is also mostly empirical and that is the reason why no theory from previous research has been used in this thesis. I have not been able to find any appropriate theory to use.

The variables used in the study are chosen based on the seven identified key areas. The best possible proxies for the seven external factors are presented in the end of the subsections and the proxies are used as explanatory variables in the statistical study.[11] As mentioned earlier, number of points in Allsvenskan is used as the explained variable, the measure of football success in Sweden.

It is of course important to evaluate the municipalities based on the characteristics of the municipalities. Nevertheless, it is important to take into account that the catchment area of a football club today can be expected to reach longer than the border of the municipality. As a proxy for the surroundings of a municipality, that can be thought of as the municipality's secondary catchment area, counties are used.[12] A county is generally so small that it represents the catchment area for a football club, both in terms of spectators and in terms of football players.

The units of the variables are chosen so that the coefficients for the variables will be similar.

2.2 History[13]

In the book "Soccernomics" from 2009, the authors Simon Kuper and Stefan Szymanski present a model of how the most important external factors affect the performance of a country's national football team. They use Russell Gerrard's soccer database of all national

[10] See the "Current state of knowledge and variable selection" section. The ways that the seven factors are expected to affect the football performance are also discussed in that section.
[11] See "Variable creation" in the "Appendix" section for a more detailed description of how the variables are constructed and from where the data used is obtained.
[12] "Län" is the Swedish word for county.
[13] The history is of course influenced by the internal factors that the clubs from a municipality have possessed in previous years. Nevertheless, the history is external for the year in focus, since the clubs cannot affect the history any more.

team results in the history and run a multiple regression (Kuper & Szymanski 2009, pp. 34-36). They concentrate on the years 1980-2001 and use goal difference as the explained variable.[14] The model manages to explain just over a quarter of the variation in goal difference. One of the conclusions drawn is that experience, measured as the accumulated number of international games, is the most important factor affecting a country's football performance. Having twice as much international experience as the opponent is worth just over half a goal. The accumulated number of games is a measure of the historical football engagement in a location and even though the results presented in the book are on a country level, it is reasonable to suspect that similar results can be shown on a Swedish municipal level.

In a C-level thesis named "Svensk Fotboll", a student at Umeå University named Elias Olofsson investigates what external factors that have been affecting a team's final position in the Allsvenskan table, the years 1994-2010 (Olofsson 2011). Olofsson uses eight variables in his regression and shows that a club's result from the previous year has a big statistically significant positive impact on the club's performance. The previous year can be seen as the most recent history. To me, a club's result from the previous year is not an entirely external factor. It is strongly affected by how well the club used its external factors in the previous year and the resource utilization rate is expected to be similar the current year.

Another fact supporting the importance of a club's, and in turn a municipality's, history is that only four new teams, without prior experience from the league, have been promoted to Allsvenskan over the eleven years of this study.[15]

As mentioned, Kuper and Szymanski use number of games as a measure of the football history (Kuper & Szymanski 2009). Nevertheless, I believe that number of points is a more accurate measure. Number of points is a measure of success and not just a measure of participation. The number of points in the Marathon table in the relevant year is therefore an explanatory variable that is used in the regression.[16] The history may also be important on a county level, so the accumulated number of points in the Marathon table for other municipalities in the same county is also an explanatory variable.

To conclude, two historical factors are used as explanatory variables:

- The municipality's number of points in the Marathon table.
- The county's number of points in the Marathon table.

2.3 Population

That international experience is important is not the only important conclusion drawn in "Soccernomics". The book also establishes that population is an important explanatory factor

[14] If a game ends 3-1, the goal difference is 2.
[15] IF Brommapojkarna, Syrianska FC, Assyriska FF, and Enköpings SK were identified by comparing the Marathon table and the tables in all the years of this study.
[16] The Marathon table is a table with the accumulated number of points for every team that has ever participated in Allsvenskan.

for the success of a country on the football pitch (Kuper & Szymanski 2009, p. 35). Having twice the opponent's population is worth about one tenth of a goal. The population density is also important. Western Europe has had a high population density and therefore new ideas have been spread fastest there (Kuper & Szymanski 2009, pp. 23-32). These network effects have been the fundament for development in terms of football, as well as for the scientific revolution.

Michael A. Leeds and Eva Marikova Leeds provide further support for the idea that population affects the football performance. In an article, named "International Soccer Success and National Institutions", they conclude that population is one of the most important factors affecting football performance and that the effect is diminishing for larger populations (Leeds & Marikova Leeds 2009, p. 370).

Finally, Olofsson includes population as one of the explanatory variables in his study (Olofsson 2011). He is unable to show that population has a statistically significant impact and that may be because he also includes average attendance as an explanatory variable. The average attendance is probably quite strongly correlated to the population size. Martin Stolt and Carl Waldenor, from Stockholm School of Economics, have shown this in a bachelor's thesis (Stolt & Waldenor 2010, p. 32).

Four relevant population factors are used as explanatory variables:

- The municipality's population.
- The county's population.
- The municipality's population density.[17]
- The county's population density.[18]

2.4 Wealth

Another important conclusion in "Soccernomics" is that the Gross Domestic Product, the GDP, per capita has a statistically significant positive impact on a country's football performance (Kuper & Szymanski 2009, p. 35). Having twice the opponent's GDP per capita is worth about one tenth of a goal. Michael A. Leeds and Eva Marikova Leeds also conclude that the GDP per capita is one of the most important factors affecting football performance (Leeds & Marikova Leeds 2009, p. 387).

How comes that wealth may affect the football performance? Wealth means more money that can be invested. More investments should generate a better outcome on the football pitch and the highest investments are likely to be observed in municipalities with a high productivity.

[17] It is easier for a club to reach both potential spectators and potential players if the distances are short. It is also easier to obtain the network effects mentioned earlier in this section.

[18] It is easier for a club to reach both potential spectators and potential players if the distances are short. It is also easier to obtain the network effects mentioned earlier in this section.

Hoffmann, Ging, and Ramasamy have shown that wealth has a diminishing positive impact on a country's football performance, up to a certain level where the impact starts to become negative (Hoffmann, Chew Ging & Ramasamy 2002, p. 266). Maybe Sweden has surpassed this level, so it is not obvious that wealth has a positive impact in the study undertaken in this thesis. Elias Olofsson's study, mentioned earlier, shows a statistically significant negative impact of wealth in the municipality (Olofsson 2011). This gives rise to the idea that a large working class population may affect the performance positively.

Gross Regional Product, GRP, is a widely accepted measure of economic wealth. Harold T. Moody and Frank W. Puffer established the concept in 1969 (Moody & Puffer 1969, p. 391) and a lot of research has been based on this measure since then. GRP is the regional equivalent to the more known national measure Gross Domestic Product, GDP, and is the market value of all final goods and services produced in a region during a specific year. Information regarding GRP is unfortunately not available on a municipality level for more than one of the eleven years, so that is why this explanatory variable only is based on county.

To conclude, one wealth factor is used as an explanatory variable:

- The county's GRP.

2.5 Working class

Most of the successful football players stem from the working class. This is a fact in England (Kuper & Szymanski 2009, pp. 18-23) and the same phenomenon can be noticed in other countries, such as Sweden. Even though the Swedish football has not been as class exclusive as the English, most Swedish football players have a working class background (Johansson 2009, p. 2). It is not for nothing that a film about IFK Göteborg's team from the middle of the 1980s, just before Swedish footballers became full time professionals, is named "Fotbollens sista proletärer" (The last proletarians of football) (Fotbollens sista proletärer 2011). The movie is about an IFK Göteborg team that was successful in the European competitions despite the fact that their opponents had full time employed players, while Göteborg's team was built upon players working as cooks, plumbers, truck drivers, etcetera.

The fact that Swedish football has had a working class touch over the years is not strange, since it was Scottish workers that took the football to Sweden (Lundberg 2010, pp. 52-53). It is thus from a foundation in the working class that the Swedish football has developed and therefore it is natural that the sport has kept a working class touch.

Today, most of the players in a team in Allsvenskan are bought from other Swedish clubs or from abroad, but it is still valuable for a club to be able to produce its own players. If a large working class population is an advantage in this area, the working class factor may still have an impact on the football performance of a Swedish municipality.

In Sweden, there is no official definition of working class and accordingly it is impossible to find any statistics showing the working class percentage in a municipality. My assessment is

that one of the most important factors in forming a person's "class identity" is the education.[19] Therefore education is a good measure to detect working class municipalities.[20] The educational level in the county may also make an impact on the football performance of a municipality, so one variable is constructed based on municipality and one is constructed based on county.

To conclude, two working class factors are used as explanatory variables:

- The municipality's educational level.
- The county's educational level.

2.6 Competition

Competition is a highly disputed topic. Some scholars believe that competition is the key for development. Others see competition as a negative factor that should be avoided. In terms of football, competition can be both beneficial and harmful for a municipality. There is a trade-off between spurring competition and competition for the same spectators, sponsors, and athletes. Competition is positive if it makes the municipal cake, in terms of spectators, sponsorships, and good athletes, bigger. If, on the other hand, resources are spent only to change the distribution of the cake, competition is harmful for the sports clubs in a municipality.

In a paper named "Status competition and performance", the authors Christoph H. Loch, Bernardo A. Huberman, and Suzanne Stout study the competition for status in a work group (Loch, Huberman & Stout 2000, p. 35). The results show that competition based on merit can push members to work hard. If competition, on the other hand, is based on political manoeuvring, the result may be negative. We can expect to see a similar pattern for competition in football. If the different clubs in a competitive area work hard, with for example marketing, the competition may result in higher attendances and better performance for all the clubs. If competition, on the other hand, is based on which club that can invite the municipal management to the fanciest dinners to increase the sponsorship, the overall sports performance in the municipality can not be expected to be influenced positively.

Elias Olofsson investigates the competition effect in a football context (Olofsson 2011). He shows that competition from other football teams in the same municipality may affect a club's performance positively.

Competition from other sports may also affect the football performance of a municipally. One interesting study from the United States sheds light on the competition between sports (Pelnar 2009, p. 2). Gregory J. Pelnar shows that having one or more other big sports in the same metropolitan area does not have a statistically significant impact on the ticket prices.

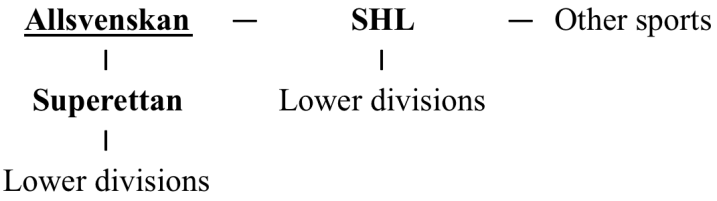
[19] The working class is, in general, characterized by a low educational level.
[20] Since GRP per capita in the county is another variable, it would not have been good to use an income measure as the variable for working class.

Furthermore, the report offers at best weak evidence that cooperation between sport franchises affect the ticket prices.

There is of course a competition between teams in Allsvenskan. The number of teams that a municipality has in Allsvenskan is rather a part of the output than an external factor. There is, however, also a competition within a county. In Stockholm County, AIK from Solna Municipality compete with Djurgårdens IF and Hammarby IF from Stockholm Municipality over fans, sponsors, and talented players. Accordingly, the number of teams in Allsvenskan from the county, but not from the municipality, is an explanatory variable used in this study.

Teams from lower divisions also contribute to the competition. To limit the scope, only teams from Superettan will be considered in this thesis. The explanatory variable is constructed on a county level and there is an obvious reason why no municipal competition effect from teams in Superettan is used in this study. If a municipality has only one good football team, that is sometimes in Allsvenskan and sometimes in Superettan, there is of course a negative correlation between the number teams in Superettan and the performance in Allsvenskan. The coefficient for the competition proxy on a municipality level would have been seriously downward biased. The municipal effect must therefore be excluded from the county variable.

Competition from other sports is of course also present in a municipality. Football and ice hockey are, by far, the biggest sports in Sweden when it comes to spectators (Sveriges Television 2014). The second divisions in both football and ice hockey have higher average attendances than the first divisions in all other sports. Therefore, the municipality's number of teams in the highest ice hockey division, SHL, has been selected as one of the explanatory variables. More sports could have been included in the research, but to limit the scope only ice hockey has been selected. The number of teams in SHL may also be interesting on a county level, so one of the explanatory variables constructed is based on municipality and one is based on county.



The closest competitors for a municipality's football clubs are clubs in Allsvenskan, clubs in Superettan and clubs in SHL.

To conclude, four competition factors are used as explanatory variables:

- The number of competing teams in Allsvenskan, from the county.
- The number of competing teams in Superettan, from the county.
- The number of competing teams in SHL, from the municipality.
- The number of competing teams in SHL, from the county.

2.7 Climate

Hoffmann, Ging, and Ramasamy have shown that the relationship between a country's football performance and the deviation of the yearly average temperature from 14°C, constitutes an inverted U-shape (Hoffmann, Chew Ging & Ramasamy 2002, p. 266). This means that the temperature has a positive impact on a country's football performance, up to a level of 14°C, where the impact starts to become negative.

"Soccernomics" reveals a number of other important conclusions beside those from the main model with international experience, population, and GDP per capita as explanatory variables. The climate is, according to the authors, another important external factor (Kuper & Szymanski 2009, p. 25). Western Europe has a mild and rainy climate, which creates a fertile land where many people want to live. The climate has accordingly led to the effects of a high population density, mentioned under the "Population" headline. Thus, both the climate and the population density may affect the success in terms of football, and these two external factors may be correlated.

In a working paper named "Does the Environment Still Matter? Daily Temperature and Income in the United States" the authors Tatyana Deryugina and Solomon M. Hsiang conclude that the productivity in the United States declines by 1.7% for each 1°C over 15°C (Deryugina & Hsiang 2014, p. 1). A weekday above 30°C costs an average American county $20 per person. It is reasonable to believe that the temperature has an impact on the productivity in Sweden too, but here it is maybe low temperatures that cost society money. It is also reasonable to believe that an outdoor activity such as football is affected even more.

Martin Stolt and Carl Waldenor from Stockholm School of Economics have shown, in a bachelor's thesis, that the temperature prior to kick off affects the attendance of a football game in Sweden positively (Stolt & Waldenor 2010, p. 32).

Temperature is perhaps the most important climate factor and an obvious explanatory variable in this study. The temperature determines if it is possible to play football outdoors or not. Unfortunately, data is only available for 84/290 municipalities. This may result in a serious selection bias. Therefore, the temperature on a municipality level cannot be used as an explanatory variable, so the variable is constructed based on county. The precipitation is another important climate factor and this explanatory variable is also constructed based on county, for the same reason. It would have been interesting to use the average temperature and precipitation in different seasons, but to limit the scope only the yearly average temperature and precipitation is used.

To conclude, two climate factors are used as explanatory variables:

- The county's yearly average temperature.
- The county's yearly precipitation.

2.8 Politics

Michael A. Leeds and Eva Marikova Leeds emphasize the importance of the political structure (Leeds & Marikova Leeds 2009, p. 387). They conclude that countries with communist regimes perform worse on the football pitch than other countries. This result is consistent with Kuper and Szymanski's reasoning in "Soccernomics" (Kuper & Szymanski 2009, pp. 23-32). Kuper and Szymanski believe that it is bad for the football development to be isolated, since knowledge is built in networks. The differences between the influential political forces in Sweden are not as big as the difference between a communist regime and another regime and the effect of the political policy on a municipal level is not total, since the national and the county policies have an impact as well. Nevertheless, it is reasonable to believe that different political policies can affect the football performances of Swedish municipalities.

Hoffmann, Ging, and Ramasamy also illustrate the big influence that politics has on sports. They conclude that the government's sports policy affects a country's football performance (Hoffmann, Chew Ging & Ramasamy 2002, p. 271).

Elias Olofsson uses a club's municipality's spending on the post "Leisure" as a measure of the municipality's financial support to the club in Allsvenskan (Olofsson 2011). That post, however, includes a lot of other spending than support to the football clubs. Another problem with the variable is that it does not include indirect municipal support. This, combined with the fact that the variable has no significant impact in Olofsson's research, drives me to conclude that this variable should not be considered in my thesis. It is, nevertheless, reasonable to believe that political decisions influence the chances for a club, or in this thesis a municipality, to reach success on the football pitch. The political governance may affect both the municipal support and other more indirect factors, and there is a significant difference in the support between different municipalities. "Sportbladet" has disclosed that the differences in municipal aids for different clubs in SHL, the highest Swedish ice hockey division, are up to SEK 19 million (Käck & Andersson 2015). As mentioned in the "Introduction" section, society affects football.

The most accurate, feasible measures of the political policy are the percentages of the votes that different political parties have in a municipality. If the percentages of the votes for all parties are used as explanatory variables, the political variables together always sum up to 100% for a municipality. The explanatory variables are accordingly, per definition, strongly correlated if all parties are used as explanatory variables. If one party decreases its share of the votes, other parties must, per definition, increase their shares of the votes. Thus, it is theoretically impossible to hold all other variables fixed when the effect of one of the political variables is investigated. Changes in one of the political explanatory variables are not independent of changes in the other political variables. To avoid such a situation, with a big multicollinearity problem, only one political variable is constructed.[21]

[21] Multicollinearity leads to worse estimates of all coefficients and makes the political variables harder to interpret.

The political variable constructed is based on the percentage of the votes for Socialdemokraterna and Vänsterpartiet, the socialist parties, in the municipal election. The municipal aids are likely to be higher in municipalities where socialists are influential and other more indirect factors may also differ depending on the support for socialist parties. Therefore I believe that the total percentage of the votes for Socialdemokraterna and Vänsterpartiet in the municipal election is a good proxy for the political policy, in this context.

To conclude, one political factor is used as an explanatory variable:

- The percentage of the votes for Socialdemokraterna and Vänsterpartiet in the municipal election.

3 Method[22]

To examine what external factors affect a Swedish municipality's chances of being successful in football, and in which ways, a statistical analysis is done. The regression run is based on panel data for the years 2003-2012.[23] Panel data is data in which multiple cases, in this case municipalities, are observed at two or more time periods (Princeton University Library 2007). It is reasonable to use panel data in this case because points in Allsvenskan, the explained variable, are generated on a yearly basis for each municipality and most statistics is published yearly.

To be able to perform a panel data regression, one of three basic models dealing with panel data must be chosen (Princeton University Library 2007):

- The First Differences model is used if the omitted variables vary over time but are the same in all municipalities. This is unlikely the case here.
- The Fixed Effects model is used if the omitted variables are time invariant, but vary between the municipalities.[24]
- The Random Effects model is used if some of the omitted variables are constant over time but vary between the municipalities, while others are fixed between municipalities but vary over time.

Since most of the omitted variables are expected to vary little over time and more between the municipalities, it is reasonable to believe that the Fixed Effects model is the right model to use. One of the biggest advantages with the Fixed Effects model, in this context, is that the time invariant omitted variables are allowed to be correlated with the explanatory variables. One disadvantage is that only the effects of factors that change over time are estimated. To make sure that the Fixed Effects model can be used, a Hausman test is run.[25]

[22] See "Tools used" in the "Appendix" section.
[23] 2013 is used to evaluate the accuracy of the estimation constructed by the regression and therefore 2013 is not included in the regression.
[24] Time invariant means invariant both in terms of values and in terms of their effects on the number of points a municipality takes in Allsvenskan.
[25] See "Model selection" in the "Tests of data" section in the "Appendix" for the result of this test.

The Fixed Effects model starts from an unobserved effects model. An unobserved effects model, dealing with panel data, can be written as follows:

$y_{it}=\beta_0+\beta_1*x_{it1}+\beta_2*x_{it2}+\ldots+\beta_k*x_{itk}+a_i+u_{it}$

y_{it} represents the number of points a specific municipality, i, takes in Allsvenskan in a specific year, t. t is in this case 2003, 2004, …, 2012. The x's are explanatory variables, the external factors that differ between municipalities and years. k is the number of explanatory variables. β_0 is a constant. The other β's, the coefficients, measure the impact of the different explanatory variables on y_{it}. a_i is a fixed effect. It captures all unobserved, time constant factors that affect the number of points a specific municipality takes in Allsvenskan. u_{it} is the idiosyncratic error and represents unobserved factors that change over time and affect the number of points a specific municipality takes in Allsvenskan, in a specific year. For each specific municipality, i, the average number of points per year can be written as follows:

$\bar{y}_i=\beta_0+\beta_1*\bar{x}_{i1}+\beta_2*\bar{x}_{i2}+\ldots+\beta_k*\bar{x}_{ik}+a_i+\bar{u}_i$

To obtain the Fixed Effects model, the average over time is subtracted from the unobserved effects model:

$y_{it}-\bar{y}_i=$

$(\beta_0+\beta_1*x_{it1}+\beta_2*x_{it2}+\ldots+\beta_k*x_{itk}+a_i+u_{it})-(\beta_0+\beta_1*\bar{x}_{i1}+\beta_2*\bar{x}_{i2}+\ldots+\beta_k*\bar{x}_{ik}+a_i+\bar{u}_i)=$

$(\beta_0-\beta_0)+\beta_1(x_{it1}-\bar{x}_{i1})+\beta_2(x_{it2}-\bar{x}_{i2})+\ldots+\beta_k(x_{itk}-\bar{x}_{ik})+(a_i-a_i)+(u_{it}-\bar{u}_i)=$

$\beta_1(x_{it1}-\bar{x}_{i1})+\beta_2(x_{it2}-\bar{x}_{i2})+\ldots+\beta_k(x_{itk}-\bar{x}_{ik})+(u_{it}-\bar{u}_i)=$

$\beta_1*\ddot{x}_{it1}+\beta_2*\ddot{x}_{it2}+\ldots+\beta_k*\ddot{x}_{itk}+\ddot{u}_{it}$

That equation is, in turn, estimated by pooled Ordinary Least Squares. Translated into everyday language, the coefficients that the model estimates show the impact of a change in each explanatory variable compared to the variable's average 2003-2012.[26] To deal with autocorrelation and heteroskedasticity, the regression is run with robust standard errors.[27]

4 Data[28]

The data used in this study is a set of panel data for all Swedish municipalities over a period of 11 years, from 2003 to 2013. 2013 is used to evaluate the prediction ability of the model, so 2013 is neither included in the regression nor in the descriptive statistics in the end of this section. Since all Swedish municipalities are included, there is no selection bias. The sample in the study equals the entire population of municipalities. There are no missing observations in the dataset, so missing observations is no issue. In general, the data is good, but there are of course some problems. One potential problem with the data is that a maximum of 16/290≈5.5% of the municipalities may be represented in Allsvenskan, each year.[29] Furthermore, a maximum of three municipalities may be replaced, each year. Only 20 municipalities have been represented in Allsvenskan in the years of this study, so the regression is really dependent on the characteristics of these municipalities. There is a similar

[26] The deviation from the mean is another way to phrase the same thing.

[27] See "Autocorrelation and heteroskedasticity" under the headline "Tests of data" in the "Appendix" section for a test of the assumption that autocorrelation and heteroskedasticity prevails.

[28] See "Variable creation" in the "Appendix" section for a more detailed description of how the variables are constructed and from where the data used is obtained.

[29] 2003-2007 a maximum of 14 municipalities could be represented in Allsvenskan, each year.

problem regarding Marathon table points in the municipality. Only 36/290≈12.4% of the municipalities have ever been represented in Allsvenskan.

For population, the highest deviations from the mean value are expected to be seen in big municipalities and counties, so the coefficients of the four population variables will be dependent on the football performance of a few well populated municipalities and counties.

The fact that a Fixed Effects model, based on deviations from the mean, is used makes it unlikely that the climate factors will have statistically significant impacts on the football performance of a municipality. Changes in temperature and precipitation are expected to be similar across the country and therefore no statistically significant correlation with the football performance is expected.

Table 1: Descriptive statistics

Category	Variable	Explanation	Obs	Mean	Std. Dev.	Min	Max
Football performance in Sweden	points	Number of points in Allsvenskan	2900	2.13	11.61	0.00	165.00
History	maramun	Number of points in the Marathon table (in hundred points)	2900	1.15	5.82	0.00	73.80
	coumara	Number of points in the Marathon table in the county, but not in the municipality (in hundred points)	2900	31.37	35.01	0.00	98.22
Population	popmun	Population (in thousand persons)	2900	31.86	62.41	2.42	881.24
	coupop	Population in the county, but not in the municipality (in thousand persons)	2900	714.07	637.51	0.00	2117.56
	popdenmun	Population density (in hundred persons/km^2)	2900	1.32	4.52	0.00	47.08
	popdencou	Population density in the county (in hundred persons/km^2)	2900	0.65	0.81	0.03	3.26
Wealth	grp	GRP per capita in the county (in SEK hundred thousand)	2900	3.14	0.59	2.20	5.48
Working class	colmun	Percentage of population with more than high school studies	2900	23.23	8.00	11.96	61.32
	colcou	Percentage of population in the county with more than high school studies	2900	29.04	4.81	20.74	42.08
Competition	couas	Number of teams in Allsvenskan from the county, but not from the municipality	2900	1.39	1.62	0.00	5.00
	couse	Number of teams in Superettan from the county, but not from the municipality	2900	1.15	1.20	0.00	4.00
	shlmun	Number of teams in SHL	2900	0.04	0.20	0.00	1.00
	coushl	Number of teams in SHL from the county, but not from the municipality	2900	0.72	0.62	0.00	2.50
Climate	temp	Average yearly temperature in the county (in °C)	2900	6.14	2.11	-0.88	9.20
	prec	Yearly precipitation in the county (in meter)	2900	0.68	0.14	0.43	1.18
Politics	soc	Percentage of the votes for Socialdemokraterna and Vänsterpartiet in the municipal election	2900	43.66	11.53	7.20	82.30

5 Results and discussion

5.1 Regression results

The table below shows the regression results from the Fixed Effects regression, with points in Allsvenskan as the explained variable.[30] The standard errors are robust and the statistically significant variables are marked in bold. The different explanatory variables are commented, by category, in the coming subsections.

Table 2: Regression results

Category	Variable	Coef.	Std.Err.	t	P>|t|	[95% Conf. Interval]	
History	**maramun**	**3.921346**	**1.565989**	**2.50**	**0.013**	**0.8391561**	**7.003535**
	coumara	-0.1841688	0.1166607	-1.58	0.116	-0.4137811	0.0454435
Population	**popmun**	**-0.6954266**	**0.2224796**	**-3.13**	**0.002**	**-1.133312**	**-0.2575409**
	coupop	0.027424	0.0167055	1.64	0.102	-0.0054558	0.0603038
	popdenmun	-0.0523003	0.8660447	-0.06	0.952	-1.756855	1.652255
	popdencou	-3.114816	8.492069	-0.37	0.714	-19.82896	13.59933
Wealth	grp	-1.053399	0.7758714	-1.36	0.176	-2.580474	0.473676
Working class	**colmun**	**0.808799**	**0.2540485**	**3.18**	**0.002**	**0.3087792**	**1.308819**
	colcou	-0.3534688	0.2611248	-1.35	0.177	-0.8674163	0.1604787
Competition	**couas**	**-0.3832469**	**0.2184634**	**-1.75**	**0.080**	**-0.8132279**	**0.0467341**
	couse	-0.0671368	0.1490876	-0.45	0.653	-0.360572	0.2262985
	shlmun	-2.23543	2.556516	-0.87	0.383	-7.267182	2.796321
	coushl	0.076965	0.2943415	0.26	0.794	-0.5023598	0.6562898
Climate	temp	0.0575456	0.0888304	0.65	0.518	-0.117291	0.2323822
	prec	0.3128025	0.8718495	0.36	0.720	-1.403177	2.028782
Politics	soc	-0.0166705	0.0317741	-0.52	0.600	-0.0792084	0.0458674
	_cons	3.662209	7.949232	0.46	0.645	-11.98352	19.30794

5.2 History

Increases in the number of points in the Marathon table for a municipality have a statistically significant positive impact on the performance in Allsvenskan, on any significance level over 1.3%. If the deviation from the mean increases by hundred Marathon table points, approximately 3.9 additional points are expected to be generated in Allsvenskan that year.[31] It takes on average 25.5 additional points in the Marathon table to generate one additional point in Allsvenskan.[32] For the observation with the highest number of points in the Marathon table 2013, in comparison to the municipality's average number of Marathon table points 2003-2012[33], the additional Marathon table points are expected to generate 25.0 additional points in Allsvenskan 2013[34]. The data supports that the most recent football history in a municipality

[30] See "Explanation of table" in the "Appendix" section for an explanation of the table.
[31] Hundred Marathon table points means an increase by one unit in the variable.
[32] 1/(Coef./100)=1/(3.921346/100)≈25.5.
[33] Göteborg Municipality.
[34] (The municipality's number of points in the Marathon table 2013−the municipality's average number of Marathon table points 2003–2012)*Coef.=(74.88−68.50)*3.921346≈25.0.

has an impact on the municipality's success in Allsvenskan. This result was expected, based on previous research.

The total number of points in the Marathon table will, per definition, increase yearly. Either a municipality's number of points in the Marathon table increases, or it remains stagnant. Therefore, this variable has not a negative impact on any municipality. It is reasonable to believe that the value of the coefficient will decrease over time, if the impact of the history remains stagnant and the average is still based on the years 2003-2012. This tendency makes the coefficient for a municipality's points in the Marathon table inappropriate to use in the long run.

The change in the number of points in the Marathon table, for all other municipalities in the county, has not a statistically significant impact on the number of points a municipality takes in Allsvenskan, on any reasonable significance level. The standard error is big, in comparison to the coefficient. Maybe the football culture associated with historical success is local and not as widespread as a county.

5.3 Population

Population increases in a municipality have a statistically significant negative impact on the municipality's performance in Allsvenskan, on any significance level over 0.2%.[35] If the deviation from the mean increases by thousand people, approximately −0.7 additional points are expected to be generated in Allsvenskan that year.[36] For each 1438 people a municipality adds to its population, the municipality is expected to lose one point in Allsvenskan.[37] For the observation with the highest population 2013, in comparison to the municipality's average population 2003-2012[38], the 86898 additional people are expected to generate 60.4 minus points in Allsvenskan 2013[39]. For the observation with the lowest population 2013, in comparison to the municipality's average population 2003-2012[40], the population decrease of 1083 people is expected to generate almost 0.8 points in Allsvenskan 2013[41].

That population increases have a negative impact on the football performance of a municipality was unexpected, since previous research suggested a positive impact. Since the three municipalities with the biggest population change[42] account for over 31% of the change[43], in absolute numbers, the coefficient has been strongly influenced by a few observations. Stockholm Municipality is responsible for over 18% of the change in population

[35] Decreases in the population have positive effects.
[36] Thousand people means an increase by one unit in the variable.
[37] −1/(Coef./1000)=−1/((−0.6954266)/1000)≈1438.
[38] Stockholm Municipality.
[39] (The municipality's population 2013−the municipality's average population 2003–2012)*Coef.=(897.700−810.802)*(−0.6954266)≈−60.4.
[40] Kramfors Municipality.
[41] (The municipality's population 2013−the municipality's average population 2003–2012)*Coef.=(18.450−19.533)*(−0.6954266)≈0.8.
[42] Stockholm Municipality, Göteborg Municipality, and Malmö Municipality.
[43] (86.898+33.730+27.531)/474.401≈31%.

2013 in comparison to the mean 2003-2012.[44] The coefficient takes a positive, but statistically insignificant, value if either Stockholm Municipality or all the three municipalities with the biggest population changes are removed. Deteriorating performance for Stockholm Municipality is responsible for the negative coefficient. What can be stated with certainty is that data does not support a positive impact of population on the performance in Allsvenskan.

Changes in the population in the rest of the county, the population density in the municipality, and the population density in the county have not a statistically significant impact on the number of points a municipality takes in Allsvenskan, on any reasonable significance level.

5.4 Wealth

Changes in the GRP per capita in the county have not a statistically significant impact on the number of points a municipality takes in Allsvenskan, on any reasonable significance level. The data does not support a positive impact of wealth on the football performance of a Swedish municipality. Maybe the positive effect of a high GRP per capita, that more investments are undertaken, is outweighed by the negative correlation of a high GRP per capita with a large working class population.

5.5 Working class

An increase in the educational level in a municipality has a statistically significant positive impact on the football performance of the municipality, on any significance level over 0.2%. If the deviation from the mean increases by one percentage point, approximately 0.8 additional points are expected to be generated in Allsvenskan that year.[45] It takes on average a 1.2 percentage points increase in the educational level to generate one additional point in Allsvenskan.[46] For the observation with the highest educational level 2013, in comparison to the municipality's average educational level 2003-2012[47], the increase of 5.70 percentage points is expected to generate 4.6 points in Allsvenskan 2013[48]. For the observation with the lowest educational level 2013, in comparison to the municipality's average population 2003-2012[49], the increase of 0.88 percentage points is expected to generate 0.7 points in Allsvenskan 2013[50]. This effect is accordingly a lot smaller than the effect of the football history in the municipality.

To conclude, the data supports that increases in the educational level in the municipality have a positive impact on the football performance. This result was unexpected, since previous research suggested a negative impact of educational level on football success. Maybe the

[44] 86.898/474.401≈18%.
[45] One percentage point means an increase by one unit in the variable.
[46] 1/Coef.=1/0.808799≈1.2.
[47] Solna Municipality.
[48] (The municipality's educational level 2013−the municipality's average educational level 2003–2012)*Coef.=(52.78−47.08)*0.808799≈4.6.
[49] Arvidsjaur Municipality.
[50] (The municipality's educational level 2013−the municipality's average educational level 2003–2012)*Coef.=(20.72−19.84)*0.808799≈0.7.

educational level is correlated to some other factors, affecting the football performance positively. This needs to be investigated further.

Changes in the educational level in the county have not a statistically significant impact on the number of points a municipality takes in Allsvenskan, on any reasonable significance level.

5.6 Competition

To get an additional team in Allsvenskan from the same county, but from another municipality, has a statistically significant negative impact on the performance in Allsvenskan.[51] This can be concluded on a significance level of 8.0%. The actual effect is, however, really small. Especially since only Östergötland County has more than one team more in Allsvenskan 2013, than the average 2003-2012. If the deviation from the mean increases by one team in Allsvenskan, approximately −0.4 additional points are expected to be generated in Allsvenskan that year.[52] It takes on average 2.6 more teams in Allsvenskan from the county, but not from the municipality, to generate one minus point in Allsvenskan.[53] The data supports that competition from teams in Allsvenskan from the same county, but another municipality, has a negative impact on the football performance of a municipality.

Changes in the number of teams in Superettan from the county, but not from the municipality, have not a statistically significant impact on the number of points a municipality takes in Allsvenskan, on any reasonable significance level. The same is the case for changes in the number of teams in SHL from the municipality and in the number of teams in SHL from the county, but not from the municipality. Clubs in other sports or lower divisions do not seem to be as big competitors for a municipality's football clubs, as other football clubs in the highest division.

5.7 Climate

Neither the change in the yearly average temperature in the county nor the change in yearly precipitation in the county has a statistically significant impact on the number of points a municipality takes in Allsvenskan, on any reasonable significance level. Given that a Fixed Effects model, based on changes in the variables, is used this result was not surprising. The climate may have an impact but changes in the climate are inconsiderable.

5.8 Politics

Changes in the percentage of the votes for Socialdemokraterna and Vänsterpartiet in the municipal elections have not a statistically significant impact on the number of points a municipality takes in Allsvenskan, on any reasonable significance level. Maybe the differences between political parties in Sweden are too small to affect the football performance of a municipality. Another possible explanation is that a political effect stems

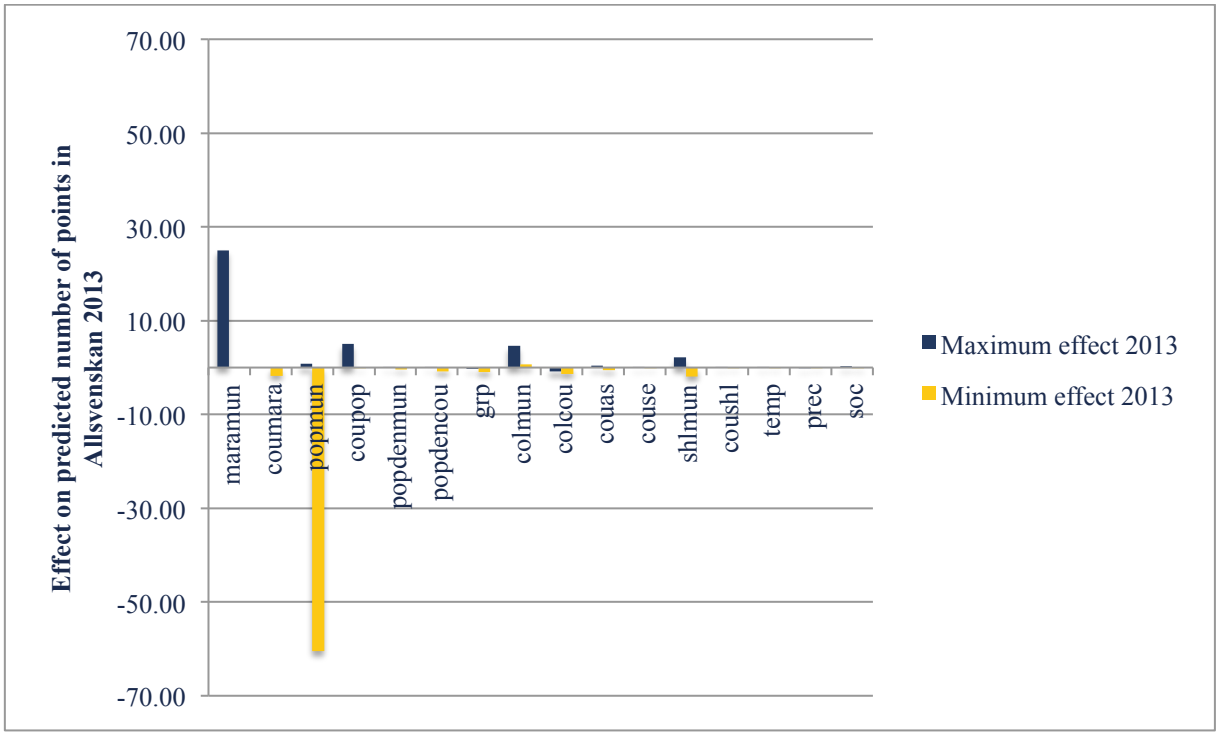
[51] If a team in Allsvenskan from the county gets relegated, there is a corresponding positive impact on the performance.
[52] One team in Allsvenskan means an increase by one unit in the variable.
[53] $-1/\text{Coef.} = -1/(-0.3832469) \approx 2.6$.

from the football interest of the politicians in the municipality, rather than from the ideology of their party.

6 Conclusions and implications

6.1 Key external factors

Diagram 1: Minimum and maximum effect of the variables 2013



The diagram shows the highest and the lowest effect on the predicted number of points in Allsvenskan 2013 that each explanatory variable has on any municipality.

4/16 chosen explanatory variables and 4/7 identified main areas have a statistically significant impact on the football performance of a Swedish municipality. Only for wealth, climate, and politics, none of the selected variables is statistically significant. Population in the municipality, however, only has a statistically significant negative impact if Stockholm is included in the sample and therefore no conclusion can be drawn regarding the population.[54]

The most important external factor is the number of points a municipality has added to the Marathon table in the most recent years. The effect is, as expected, positive. A more surprising finding is that an increase in the educational level in the municipality improves the football performance. Competition from an additional team in Allsvenskan from the same county, but not from the same municipality, has a small negative impact on the football performance of a municipality.

[54] The coefficient is, as stated earlier, positive and statistically insignificant if Stockholm is removed.

6.2 Evaluation of the municipalities

The coefficients that the regression produces are used to rank the municipalities in 2013, after their chances to achieve success in Allsvenskan. Each coefficient is multiplied by its explanatory variable's value 2013 minus the explanatory variable's average value 2003-2012.[55] After that, the average number of points 2003-2012 is added, to obtain the prediction for the municipality in 2013.[56] The number obtained is compared to the actual performance in 2013. The Fixed Effects are predicted by Stata.

Table 4: The municipalities with the highest predicted number of points 2013

Municipality	Fixed Effect	Predicted points 2013	Points 2013	Difference
Göteborg	-37.12	117.83	91	-26.83
Borås	47.22	59.80	46	-13.80
Helsingborg	26.32	56.60	49	-7.60
Kalmar	5.55	52.07	52	-0.07
Solna	119.98	50.22	58	7.78
Malmö	-41.48	45.86	63	17.14
Stockholm	-404.83	42.43	76	33.57
Halmstad	24.77	36.92	31	-5.92
Gävle	-41.01	33.96	34	0.04
Örebro	-21.27	31.65	0	-31.65
Trelleborg	9.54	24.25	0	-24.25
Sölvesborg	0.90	17.11	36	18.89
Sundsvall	-31.58	15.88	0	-15.88
Landskrona	34.08	11.01	0	-11.01
Norrköping	45.50	10.75	39	28.25
Södertälje	-23.88	9.58	14	4.42
Åtvidaberg	20.31	9.55	40	30.45

6.2.1 Prediction ability

The table above, containing the 17 municipalities with the highest predicted number of points 2013 by the regression results, contains 13/14 municipalities represented in Allsvenskan 2013.[57] From that, it can be concluded that the model predicts the actual outcome, which is of course also affected by the internal usage of the external factors, quite well.

At a first glance, the prediction ability is surprisingly high, since most municipalities have a maximum of one good football team. That fact makes them extremely vulnerable to the internal resource usage, affected by a few individuals in the only club. Big deviations from the predicted values in single years were accordingly expected and the model was expected to be

[55] This is done in Excel.
[56] $(y_{it}-\bar{y}_i)+\bar{y}_i=y_{it}$.
[57] Only Växjö Municipality is outside the list.

better at predicting long run performance. One explanation to the good prediction ability is that the Fixed Effects for each municipality are included in the predicted numbers.[58]

6.2.2 Fixed Effects

As you can see in the table in the beginning of the "Evaluation of the municipalities" section, the Fixed Effects are quite big in comparison to the predicted number of points. The mean absolute value of the Fixed Effects is 13.92 and the mean absolute value of the predicted number of points 2013 after the Fixed Effects have been subtracted is 14.30. From that it can be concluded that the Fixed Effects are responsible for around 49% of the predicted number of points in Allsvenskan[59], while the explanatory variables are responsible for approximately 51%[60].

Table 5: The municipalities with the highest Fixed Effects

Municipality	Fixed Effect
Solna	119.98
Borås	47.22
Norrköping	45.50

This table shows the three municipalities with the highest Fixed Effects. The deviations from the means for all the explanatory variables in this study have not been enough to explain this part of the predictions. These three municipalities should be investigated further to examine if they have some common characteristics and to find the components of their Fixed Effects. The fact that Solna and Stockholm have grown together to a great extent may explain the big Fixed Effect for Solna Municipality. The urban areas in most municipalities are longer away from other municipalities' urban areas. This fact may also explain a part of the negative Fixed Effect for Stockholm Municipality.

Table 6: The municipalities with the lowest Fixed Effects

Municipality	Fixed Effect
Stockholm	-404.83
Uppsala	-102.24
Linköping	-76.44
Västerås	-70.47
Jönköping	-56.66
Umeå	-48.19

This table shows the six municipalities with the lowest Fixed Effects. The deviations from the means for all the explanatory variables in this study have not been enough to explain this part

[58] $(y_{it}-\bar{y}_i)+\bar{y}_i=y_{it}=\beta_0+\beta_1*x_{it1}+\beta_2*x_{it2}+\ldots+\beta_k*x_{itk}+a_i+u_{it}$.
[59] 13.92/(13.92+14.30)≈49%.
[60] 14.30/(13.92+14.30)≈51%.

of the predictions. These six municipalities should be investigated further to examine if they have some common characteristics and to find the components of their Fixed Effects. It is noteworthy that four of these municipalities are known for their Universities[61] and that four of them have cathedrals[62].

6.2.3 Expected municipalities in Allsvenskan

In 2013, 14 municipalities were represented in Allsvenskan.[63] The 14 first municipalities in the list should have had a team in Allsvenskan 2013, if they had used their external factors as well as their competitors did, given that 14 municipalities are represented in Allsvenskan. Notice that 10 of the 14 municipalities with the best external factors were represented in Allsvenskan 2013. The exceptions were Örebro Municipality, Trelleborg Municipality, Sundsvall Municipality, and Lanskrona Municipality. These municipalities are expected to improve in the long run and both Sundsvall Municipality and Örebro Municipality have teams in Allsvenskan today, 2015. For Norrköping Municipality, Södertälje Municipality, Åtvidaberg Municipality, and Växjö Municipality, the condition is the opposite. They were represented in Allsvenskan 2013 because of good internal usage of the external factors. Unsurprisingly, Växjö Municipality and Södertälje Municipality (Swedish Football Association 2013) have seen their teams getting relegated and Åtvidaberg Municipality's team is expected to struggle 2015.

6.2.4 Overachievers

Table 7: Overachievers 2013

Municipality	Fixed Effect	Predicted points 2013	Points 2013	Difference
Stockholm	-404.83	42.43	76	33.57
Åtvidaberg	20.31	9.55	40	30.45
Växjö	8.40	-0.65	28	28.65
Norrköping	45.50	10.75	39	28.25
Sölvesborg	0.90	17.11	36	18.89
Malmö	-41.48	45.86	63	17.14

Stockholm Municipality was the biggest overachiever in 2013. The big population increase in Stockholm Municipality makes the predicted number of points low. All the municipalities in this list can expect their performance to drop in the long run, if they cannot maintain their superior internal usage of the external factors. That Växjö Municipality has a negative value as the predicted number of points 2013 shows that the model is not 100% reliable. No municipality can expect to generate minus points in Allsvenskan.

[61] Stockholm Municipality, Uppsala Municipality, Linköping Municipality, and Umeå Municipality.
[62] Stockholm Municipality, Uppsala Municipality, Linköping Municipality, and Västerås Municipality.
[63] Göteborg Municipality and Stockholm Municipality were represented by two teams.

6.2.5 Underachievers

Table 8: Underachievers 2013

Municipality	Fixed Effect	Predicted points 2013	Points 2013	Difference
Örebro	-21.27	31.65	0	-31.65
Göteborg	-37.12	117.83	91	-26.83
Trelleborg	9.54	24.25	0	-24.25
Sundsvall	-31.58	15.88	0	-15.88
Borås	47.22	59.80	46	-13.80
Landskrona	34.08	11.01	0	-11.01
Helsingborg	26.32	56.60	49	-7.60
Halmstad	24.77	36.92	31	-5.92

Örebro Municipality was the municipality without representation in Allsvenskan 2013 that had the best chances to improve. Unsurprisingly, Örebro Municipality has a team playing in Allsvenskan now, 2015, and Örebro SK ended up sixth in Allsvenskan 2014 (Swedish Football Association 2014). Trelleborg Municipality, Sundsvall Municipality, and Landskrona Municipality can also expect to be represented in Allsvenskan in the future. Sundsvall Municipality is represented in Allsvenskan 2015, but more surprisingly Trelleborg Municipality and Landskrona Municipality do not even have a team in Superettan right now (Swedish Football Association 2015). They are really underperforming.

The biggest underachiever represented in Allsvenskan 2013 was Göteborg Municipality. Despite being the best football municipality 2013, the 91 points were not enough to match what the regression predicted. There is room for a third team from Göteborg in Allsvenskan.

Borås Municipality, Helsingborg Municipality, and Halmstad Municipality were other underperformers in Allsvenskan 2013.

6.2.6 Potential for a stable team in Allsvenskan

In the years 2008-2013, when 16 teams have participated in Allsvenskan, 32 has been the average number of points needed to avoid the three relegation spots.[64] With that in hand, it is possible to conclude that only nine municipalities can expect to have a team in Allsvenskan every year. These municipalities are Göteborg Municipality, Borås Municipality, Helsingborg Municipality, Kalmar Municipality, Solna Municipality, Malmö Municipality, Stockholm Municipality, Halmstad Municipality, and Gävle Municipality. Sölvesborg Municipality, Norrköping Municipality, Södertälje Municipality, Åtvidaberg Municipality, and Växjö Municipality cannot expect their teams to be in Allsvenskan every year.[65]

[64] One of the three relegation spots means playoff against the third team in Superettan.

[65] The teams from Sölvesborg Municipality (Swedish Football Association 2014), Södertälje Municipality, and Växjö Municipality (Swedish Football Association 2013) have already been relegated.

6.2.7 Few newcomers expected

The, perhaps, most important conclusion that can be drawn by studying the list is that we cannot expect to see many municipalities without prior representation to be represented in Allsvenskan in the future. All the municipalities in the list in the beginning of this section have had teams in Allsvenskan in the sample period. The most recent history seems to be really important and the most recent history consists of the points generated in the previous years. Those years may, in turn, have been influenced by their recent history. Thus, it is possible that a small advantage in the beginning of the football history still has a big impact. In the light of this, it should be noted that no teams from north of Gävle was allowed to participate in Allsvenskan until 1953/54 (Nylin 2004, p. 16). The effects from this discrimination may still be huge. This may be the reason why only one municipality from this area, Sundsvall Municipality, is in the list. Being located north of Gävle may be a negative Fixed Effect.

6.3 Further research

Since the history of football is the most important external predictor of a municipality's chances to achieve success on the football pitch, a study similar to this should be undertaken to explore the historical success factors. External factors such as the ones used in this thesis could be used to investigate the historical success. Maybe the statistically insignificant explanatory variables today have had a big impact historically. Research in other scientific areas could also be undertaken to explore the historical success. Studying the expansion of football in Sweden could be a good idea. The fact that the Scottish workers who brought the football to Sweden arrived at Göteborg is maybe the reason why Göteborg is successful today (Lundberg 2010, pp. 52-53). How has the sport spread since then and why was Göteborg the city that the Scots arrived to? Has the presence of universities and cathedrals limited the impact of football? The potential topics for further research in this area are many. It would also be interesting to investigate if the importance of the history increases or decreases over time.

Another thing that needs to be studied further is why increases in the educational level in a municipality improve the football performance. The effects of competition within Swedish football could be investigated further, as well.

The external factors that do not have a statistically significant impact in this study may also be interesting to study further. Maybe they can be split up into different factors that have cancelled each other out in this study.

It would also be interesting to do a follow up study of this one and see how well the model predicts the performance over time, in the future, and to investigate how the coefficients change over time. It is also possible to extend this research to different countries and different sports.

Another important extension of this research would be to investigate how the output, in terms of points, is translated into utility for the municipalities.

As mentioned under the headline “Fixed Effects” in the “Evaluation of the municipalities” section, the Fixed Effects should also be investigated further. Are some parts of the Fixed Effects attributable to other external factors that could have been included as explanatory variables in a study like this?

7 References

Alsiö, M 2011, *100 år med allsvensk fotboll*, Idrottsförlaget i Västerås AB, Västerås.

Cederfelt, M 2012, *riksdagen.se*, Margareta Cederfelt, Stockholm, viewed 22 April 2015, <http://www.riksdagen.se/sv/Dokument-Lagar/Forslag/Motioner/Den-svenska-staden_H002K315/?text=true>.

Deryugina, T & Hsiang, SM 2014, 'Does the Environment Still Matter? Daily Temperature and Income in the United States', Working paper, University of Illinois Urbana-Champaign and University of California, Berkeley and NBER, The National Bureau of Economic Research.

Fotbollens sista proletärer 2011, Svenska Filmstudion.

Hoffmann, R, Chew Ging, L & Ramasamy, B 2002, 'The socio-economic determinants of international soccer performance', *Journal of Applied Economics*, vol. 5, no. 2, November 2002, pp. 253-272.

Investopedia, LLC 2015, *investopedia.com*, viewed 21 March 2015, <http://www.investopedia.com/terms/a/autocorrelation.asp>.

Investopedia, LLC 2015, *investopedia.com*, viewed 21 March 2015, <http://www.investopedia.com/terms/h/heteroskedasticity.asp>.

Johansson, L 2009, 'Idrottsvärlden och idrottsvärden', Bachelor thesis, Pedagogics, Lund University, Lund.

Käck, A & Andersson, M 2015, '"Kommunal dopning så det dånar om det"', *Aftonbladet*, 7 March 2015.

Kuper, S & Szymanski, S 2009, *Soccernomics*, Nation Books, New York.

Leeds, MA & Marikova Leeds, E 2009, 'International Soccer Success and National Institutions', *Journal of Sports Economics*, vol. 10, no. 4, August 2009, pp. 369-390.

Loch, CH, Huberman, BA & Stout, S 2000, 'Status competition and performance in work groups', *Journal of Economic Behavior & Organization*, vol. 43, no. 1, September 2000, pp. 35-55.

Lundberg, B 2010, 'Fotbollens historia', *Allt om Historia*, no. 6 2010.

Moody, HT & Puffer, FW 1969, 'A Gross Regional Product approach to regional model-building', *Economic Inquiry*, vol. 7, no. 4, December 1969, pp. 391-402.

Nylin, L 2004, *Den nödvändiga boken om Allsvenskan*, Bokförlaget Semic, Uddevalla.

Olofsson, E 2011, 'Svensk Fotboll', C-level thesis, Economics, Umeå University, Umeå University, Umeå.

Pelnar, GJ 2009, 'Competition and Cooperation Between Professional Sports Franchises: The Impact on Ticket Prices', Report, Compass Lexecon, Chicago.

Princeton University Library 2007, *dss.princeton.edu*, viewed 21 March 2015, <http://dss.princeton.edu/online_help/stats_packages/stata/panel.htm>.

Princeton University Library 2007, *dss.princeton.edu*, viewed 23 March 2015, <http://dss.princeton.edu/online_help/analysis/interpreting_regression.htm>.

Rensselaer Polytechnic Institute 2013, *homepages.rpi.edu*, viewed 22 March 2015, <http://homepages.rpi.edu/~simonk/pdf/UsefulStataCommands.pdf>.

SMHI 2009, *smhi.se*, viewed 10 March 2015, <http://www.smhi.se/polopoly_fs/1.2994!temperatur_dec-2009.pdf>.

SMHI 2015, *smhi.se*, viewed 10 March 2015, <http://www.smhi.se/klimatdata/meteorologi/temperatur/2.1240>.

Statistics Sweden 2014, *scb.se*, viewed 8 March 2015, <http://www.scb.se/NR0105/#c_undefined>.

Statistics Sweden, *scb.se*, viewed 10 March 2015, <http://www.scb.se/sv_/Hitta-statistik/Regional-statistik-och-kartor/Regionala-indelningar/Lan-och-kommuner/Lan-och-kommuner-i-kodnummerordning/>.

Statistics Sweden, *statistikdatabasen.scb.se*, viewed 8 March 2015, <http://www.statistikdatabasen.scb.se/pxweb/sv/ssd/START__BE__BE0101__BE0101C/BefArealTathetKon/?rxid=3cf87ffa->.

Statistics Sweden, *statistikdatabasen.scb.se*, viewed 8 March 2015, <http://www.statistikdatabasen.scb.se/pxweb/sv/ssd/START__UF__UF0506/Utbildning/?rxid=76fda58a-658a-40ee-973b-5d935b04c855>.

Statistics Sweden, *statistikdatabasen.scb.se*, viewed 8 March 2015, <http://www.statistikdatabasen.scb.se/pxweb/sv/ssd/START__ME__ME0104__ME0104A/ME0104T1/?rxid=83b46d4b-8aa8-4fb8-992e-0142228f62db>.

Stolt, M & Waldenor, C 2010, 'The demand for football attendance in Sweden', BSc Thesis, Department of Marketing and Strategy (DMS), Stockholm School of Economics, Stockholm.

Sveriges Television 2014, *Publikbarometern: Största serierna*, Sveriges Television, viewed 8 March 2015, <http://www.svt.se/sport/article1959393.svt/binary/SVT%20Sports%20publikundersokning-torsdag.pdf>.

Swedish Association of Local Authorities and Regions 2008, *skl.se*, viewed 8 March 2015, <http://skl.se/download/18.28eae57e14983f3bbe0899f1/1417177912356/kommuner_text.pdf>.

Swedish Football Association 2011, *svenskfotboll.se*, viewed 10 March 2015, <http://svenskfotboll.se/allsvenskan/tidigare-ar/resultat-2011/tabell-och-resultat/>.

Swedish Football Association 2012, *svenskfotboll.se*, viewed 10 March 2015, <http://svenskfotboll.se/allsvenskan/tidigare-ar/resultat-2012/tabell-och-resultat/>.

Swedish Football Association 2013, *svenskfotboll.se*, viewed 10 March 2015, <http://svenskfotboll.se/allsvenskan/tidigare-ar/resultat-2013/tabell-och-resultat/>.

Swedish Football Association 2014, 'Analys av allsvenska klubbarnas ekonomi 2013', Swedish Football Association.

Swedish Football Association 2014, *svenskfotboll.se*, viewed 10 March 2015, <http://svenskfotboll.se/allsvenskan/historik/maratontabell/>.

Swedish Football Association 2014, *svenskfotboll.se*, viewed 23 March 2015, <http://svenskfotboll.se/allsvenskan/tidigare-ar/resultat-2014/tabell-och-resultat/>.

Swedish Football Association 2015, *svenskfotboll.se*, viewed 23 March 2015, <http://svenskfotboll.se/superettan/tabell-och-resultat/>.

Swedish Football Association, *svenskfotboll.se*, viewed 12 March 2015, <http://svenskfotboll.se/superettan/tidigare-ar/>.

Swedish Ice Hockey Association 2015, *stats.swehockey.se*, viewed 10 March 2015, <http://stats.swehockey.se/ScheduleAndResults/Overview/5056>.

Swedish Ice Hockey Association, *historical.stats.swehockey.se*, viewed 10 March 2015, <http://historical.stats.swehockey.se>.

Swedish Sports Confederation 2013, '2012 Verksamhetsberättelse med årsredovisningar', Annual report, Swedish Sports Confederation, Växjö.

Swedish Sports Confederation 2014, '2013 Verksamhetsberättelse med årsredovisningar', Annual report, Swedish Sports Confederation, Växjö.

Syrianska FC, *syrianskafc.com*, viewed 10 March 2015, <http://syrianskafc.com/Arenabiljetter.aspx>.

8 Appendix

8.1 Tools used

8.1.1 Excel

All the preparation of the dataset is done in Excel. In Excel it is easy to get a good overview of the dataset and to keep control of what is done. Observations for 290 municipalities in eleven years are used, so there is a lot of data that needs to be overviewed and entered correctly.

8.1.2 Stata

The regression and all the tests are run in Stata, a program well suited for analysing large datasets. Stata is fast and easy to use and the do-file, as you can see below, offers an easy way to show what is done.

***Preparations**
cd "C:\Users\22853\Documents\Thesis"
use "Dataset2003-2012.dta", clear
xtset munnr year, yearly
findit xtserial[66]
ssc install xttest3
ssc install outreg2

***Hausman test**
xtreg points maramun coumara popmun coupop popdenmun popdencou grp colmun colcou couas couse shlmun coushl temp prec soc , fe
estimates store fixed
xtreg points maramun coumara popmun coupop popdenmun popdencou grp colmun colcou couas couse shlmun coushl temp prec soc , re
estimates store random
hausman fixed random

***Test for autocorrelation**
xtserial points maramun coumara popmun coupop popdenmun popdencou grp colmun colcou couas couse shlmun coushl temp prec soc

***Test for heteroskedasticity**
xtreg points maramun coumara popmun coupop popdenmun popdencou grp colmun colcou couas couse shlmun coushl temp prec soc , fe
xttest3

[66] Click on "st0039 from http://www.stata-journal.com/software/sj3-2" and after that on "click here to install"

***Regression**

xtreg points maramun coumara popmun coupop popdenmun popdencou grp colmun colcou couas couse shlmun coushl temp prec soc , fe robust
outreg2 vote using regressionresults.doc, replace
predict fe

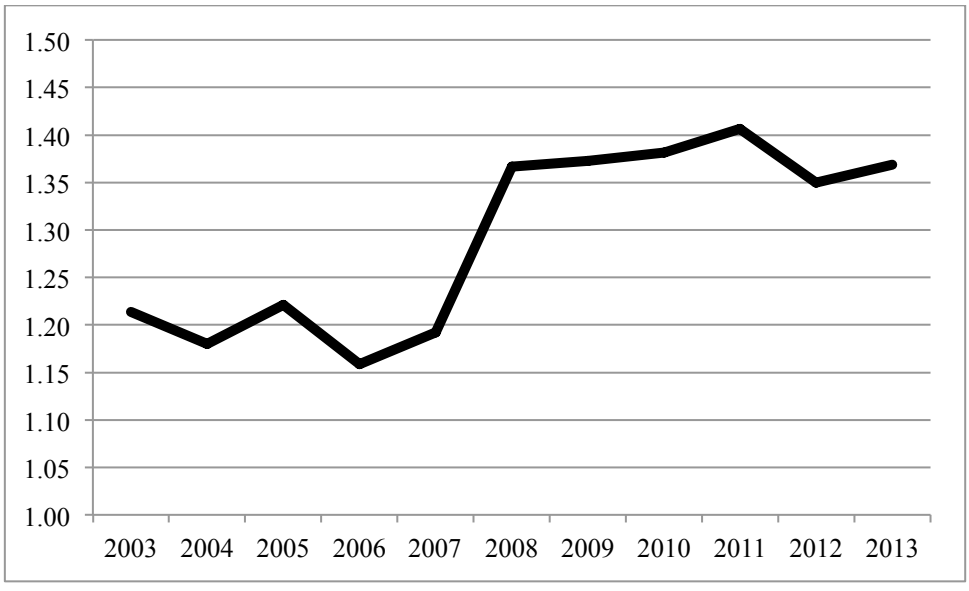
8.2 Variable creation

8.2.1 Football performance in Sweden

8.2.1.1 Points in Allsvenskan

The first step is to find information regarding points in the years, for clubs. The book "100 år med allsvensk fotboll" by Martin Alsiö is used to find the tables for the years 2003-2010 (Alsiö 2011, pp. 492-507).[67] The tables for 2011 (Swedish Football Association 2011), 2012 (Swedish Football Association 2012), and 2013 (Swedish Football Association 2013) are obtained from the website of the Swedish Football Association. The number of points for 2003-2007 have been weighted by a factor of 30/26 to make the number of points comparable to the situation from 2008 and onwards, when 30 games are played and not only 26. Despite that adjustment, the total number of points in Allsvenskan has increased significantly since the number of teams was increased to 16 2008, as the diagram below showing points per game and team enlightens. This is due to the fact that more games have had a winner and fewer games have ended up in draws.[68] The explanation to that fact is that more teams means a bigger difference in quality between the teams, and accordingly fewer draws.

Diagram 2: Average number of points per team and game in Allsvenskan



[67] It should be noted that The Swedish Football Association punished Hammarby IF by depriving three points, after lacking security in the game against Djurgårdens IF 28 August 2006.
[68] A game with a winner means three points to the winner and zero points to the loser, while a draw means one point each.

After the number of points per team each year has been determined, the next step is to connect each team to a municipality. hitta.se is used to find the locations for the clubs on a map. The map is compared to a municipality map (Swedish Association of Local Authorities and Regions 2008), to decide in which municipality each club is located. The points for all the clubs in a municipality are added to a common number of points for the municipality, in each year. Of course all the municipalities that have not had any team in Allsvenskan in this time period have zero points each year.

Table 9: Points in Allsvenskan per municipality 2003-2012

Municpality	Points 2003	Points 2004	Points 2005	Points 2006	Points 2007	Points 2008	Points 2009	Points 2010	Points 2011	Points 2012	Points 2003-2012
Göteborg	82	74	120	102	78	92	165	112	145	108	1078
Stockholm	109	78	96	83	105	77	85	65	36	37	771
Malmö	48	52	41	38	34	44	43	67	54	56	477
Borås	34	32	37	50	40	63	55	47	57	59	474
Helsingborg	38	30	39	42	35	54	43	65	63	50	459
Kalmar	0	40	43	41	48	64	50	40	44	37	407
Solna	39	25	0	49	38	45	61	35	58	55	405
Halmstad	36	50	32	27	36	41	32	35	14	0	303
Örebro	37	33	0	0	25	42	45	52	36	24	294
Gävle	0	0	31	31	34	28	39	29	41	36	269
Trelleborg	0	13	0	0	23	40	41	44	25	0	186
Sundsvall	19	34	25	0	0	22	0	0	0	29	129
Sölvesborg	0	0	0	0	0	0	0	43	40	34	117
Norrköping	0	0	0	0	0	20	0	0	34	52	106
Landskrona	32	30	30	0	0	0	0	0	0	0	92
Södertälje	0	0	14	0	0	0	0	0	28	34	76
Åtvidaberg	0	0	0	0	0	0	0	29	0	37	66
Växjö	17	0	0	19	0	0	0	0	0	0	36
Uddevalla	0	0	0	0	0	24	0	0	0	0	24
Enköping	14	0	0	0	0	0	0	0	0	0	14
Total	505	491	508	482	496	656	659	663	675	648	5783

8.2.2 History

8.2.2.1 Municipality's points in the Marathon table

The value of this variable is computed by subtracting the number of points for the remaining years through 2014 from the Marathon table after the season of 2014 (Swedish Football Association 2014), for each team.[69] After the number of points in the Marathon table for each team in each year has been determined, the next step is to connect the teams to a municipality. hitta.se is used to find the locations for the clubs on a map. The map is compared to a municipality map (Swedish Association of Local Authorities and Regions 2008), to decide in which municipality each club is located. The Marathon table points for all the clubs in a

[69] The remaining years through 2014 for example means 2004-2014 for 2003 and 2009-2014 for 2008.

municipality are added to a common number of Marathon table points for the municipality, in each year.

8.2.2.2 County's points in the Marathon table

Each municipality's number of points in the Marathon table is connected to a county by using a list from scb.se (Statistics Sweden n.d.).

8.2.3 Population[70]

8.2.3.1 Population in the municipality

The information required is obtained from Statistics Sweden (Statistics Sweden n.d.).[71] The data is rescaled and reported in thousands of people.

8.2.3.2 Population in the county

The information required is obtained from Statistics Sweden (Statistics Sweden n.d.).[72] Each municipality is connected to a county by using a list from scb.se (Statistics Sweden n.d.). To avoid having two variables containing the same information, the population in the municipality is subtracted from the county population, to form this variable. The data is rescaled and reported in thousands of people.

8.2.3.3 Population density in the municipality

The information is obtained from Statistics Sweden (Statistics Sweden n.d.).[73] The data is rescaled and reported in hundreds of people/km^2.

8.2.3.4 Population density in the county

The information is obtained from Statistics Sweden (Statistics Sweden n.d.).[74] The data is rescaled and reported in hundreds of people/km^2.

8.2.4 Wealth

8.2.4.1 GRP per capita in the county

The information is obtained from Statistics Sweden (Statistics Sweden 2014).[75] The data is rescaled and reported in SEK hundred thousand.

[70] The numbers used for the different population factors are the numbers at the end of the year in focus.
[71] The boxes "Folkmängd" and the years 2003–2013 should be ticked. Under "region", "kommuner" should be selected and all of them ticked.
[72] The boxes "Folkmängd" and the years 2003–2013 should be ticked. Under "region", "län" should be selected and all of them ticked.
[73] The boxes "Invånare per kvadratkilometer" and the years 2003–2013 should be ticked. Under "region", "kommuner" should be selected and all of them ticked.
[74] The boxes "Invånare per kvadratkilometer" and the years 2003–2013 should be ticked. Under "region", "län" should be selected and all of them ticked.
[75] "Table 3" in "Regionala Räkenskaper 2000–2013" is used.

8.2.5 Working class

8.2.5.1 Educational level in the municipality

The information is obtained from "Statistikdatabasen" at the webpage of Statistics Sweden (Statistics Sweden n.d.).[76] The variable is computed as the percentage of the population with more than high school studies.

8.2.5.2 Educational level in the county

The information is obtained from "Statistikdatabasen" at the webpage of Statistics Sweden (Statistics Sweden n.d.).[77] The variable is computed as the percentage of the population with more than high school studies.

8.2.6 Competition

8.2.6.1 Teams in Allsvenskan from the county

The first step in forming this variable is to determine the participating teams in each year. The book "100 år med allsvensk fotboll" by Martin Alsiö is used to find the tables for the years 2003-2010 (Alsiö 2011, pp. 492-507). The tables for 2011 (Swedish Football Association 2011), 2012 (Swedish Football Association 2012), and 2013 (Swedish Football Association 2013) are obtained from the website of the Swedish Football Association.

After the participating teams each year have been determined, the next step is to connect each team to a municipality. hitta.se is used to find the locations for the teams on a map. The map is compared to a municipality map (Swedish Association of Local Authorities and Regions 2008), to decide in which municipality each team is located. The teams from a municipality are added to a common number for the municipality, in each year. Each municipality's number of teams is connected to a county by using a list from scb.se (Statistics Sweden n.d.). Of course all the counties that have not had any team in Allsvenskan in this time period have a value of zero for this variable.

The last step is to subtract the number of teams from the municipality, from the number of teams from the county. The reason to this subtraction is that the number of teams a municipality has in Allsvenskan is a part of the output rather than an external factor.

8.2.6.2 Teams in Superettan from the county

The tables for 2003-2013 are found on the webpage of the Swedish Football Association (Swedish Football Association n.d.). After the participating teams in Superettan in the different seasons have been determined, the next step is to use hitta.se to find the addresses for the clubs. The addresses are compared to a municipality map (Swedish Association of Local Authorities and Regions 2008), to decide in which municipality each club is located. The clubs from all municipalities within a certain county (Swedish Association of Local

[76] "Kommuner" should be selected as the "region" and all of the municipalities should be ticked. All alternatives under "utbildningsnivå" should also be ticked, as well as the years 2003–2013.

[77] "Län" should be selected as the "region" and all of the counties should be ticked. All alternatives under "utbildningsnivå" should also be ticked, as well as the years 2003–2013.

Authorities and Regions 2008) are then added, to form a number for the county in question. The value of this variable for every single municipality is the number of teams in Superettan from the county, minus the number of teams in Superettan from the municipality.

8.2.6.3 Teams in SHL from the municipality

The data for SHL 2012/13-2013/14 is obtained at stats.swehockey.se (Swedish Ice Hockey Association 2015).[78] For the seasons 2002/03-2011/12 historical.stats.swehockey.se is used (Swedish Ice Hockey Association n.d.).[79]

After the participating teams in SHL in the different seasons have been determined, the next step is to use hitta.se to find the locations for the clubs on a map. The map is compared to a municipality map (Swedish Association of Local Authorities and Regions 2008), to decide in which municipality each club is located. Since the seasons are split over two years, 50% of the number of teams in each season is attributed to each year.

8.2.6.4 Teams in SHL from the county

The data for this variable is obtained by using the data from "Teams in SHL from the municipality" and connecting every municipality from a certain county to the county in question (Swedish Association of Local Authorities and Regions 2008). The number of teams from each municipality is added to a common value for the county. To avoid having two variables containing the same information, the number of teams in SHL from the municipality is subtracted from the county number, to form this variable.

8.2.7 Climate

8.2.7.1 Average yearly temperature in the county

The yearly average temperature in a municipality is obtained by using statistics from 131 different measuring stations in the years 2003-2013 (SMHI 2015).[80] The data on a municipality level is used to construct the county average.

To connect the measuring stations to municipalities, a map of the municipalities (Swedish Association of Local Authorities and Regions 2008) and a map of the measuring stations (SMHI 2009) are compared. The averages from the measuring stations in each municipality form the municipality value. The average values from the municipalities in each county form the county values. Each municipality is connected to a county by using a list from scb.se (Statistics Sweden n.d.).

Data for each county, except Södermanland, is available to obtain in this way. The value for Södermanland County is constructed by taking the average value for the adjacent counties Stockholm, Uppsala, Västmanland, Örebro and Östergötland.

[78] The tables for the different seasons are selected from the menu to the left.
[79] The division and the years are selected from the menu to the left. Notice that SHL was named Elitserien at this point of time.
[80] The unit of temperature is °C.

8.2.7.2 Precipitation in the county

The precipitation is obtained by using statistics from 131 different measuring stations in the years 2003-2013 (SMHI 2015). The same method as for the temperature is used. The data is rescaled and reported in meters per year.

8.2.8 Politics

8.2.8.1 Percentage of the votes for Socialdemokraterna and Vänsterpartiet in the municipal election

The data is obtained from statistikdatabasen.scb.se (Statistics Sweden n.d.). Since the elections take place in September, in the end of a football season, the results from the elections are not counted for the election year, but for the coming four years.

8.3 Tests of data

8.3.1 Model selection

To make sure that the Fixed Effects model can be used, a Hausman test is run.

Table 10: Hausman test results

chi2(16)	399.59
prob>chi2	0.0000

A high chi2 value[81] and a low P value[82] ("Prob>chi2") indicate that the Fixed Effects model should be used. The result strongly supports the usage of a Fixed Effects model. Stata included all 16 explanatory variables in the test.

8.3.2 Autocorrelation and heteroskedasticity

To choose a proper regression method, it must be clarified whether autocorrelation prevails or not. Autocorrelation is a measure of the degree of similarity between a given time series and a lagged version of itself, over successive time intervals (Investopedia, LLC 2015). The observations in the dataset are expected to be autocorrelated. That so is the case is easy to see, since few things change dramatically in a municipality, from year to year. To make sure that autocorrelation prevails, a Wooldridge test for autocorrelation in panel data is carried out in Stata.

Table 11: Wooldridge test results

F(1, 289)	11.994
prob>F	0.0006

[81] A measure of the goodness of fit.
[82] The probability of obtaining a chi2 value as large as the one observed, given that all parameters except the intercept are zero.

The table above, with the result from the Wooldridge test, shows that autocorrelation is present. A high F value[83] and a low P value[84] ("Prob>F") indicate autocorrelation. On a significance level of 0.06%, the null hypothesis of no autocorrelation can be rejected. Autocorrelation prevails and must be adjusted for in the regression. Stata included 289/290 municipalities in the test, so the test is reliable.

It must also be clarified whether heteroskedasticity prevails or not. Heteroskedasticity prevails if the standard deviations of the variables, monitored over a specific amount of time, are not constant (Investopedia, LLC 2015). It is likely that heteroskedasticity will prevail. That is the case in most datasets. To check for heteroskedasticity, a modified Wald test for groupwise heteroskedasticity in Fixed Effect regression model is carried out in Stata.

Table 12: Modified Wald test results

chi2(290)	130000000
prob>chi2	0.0000

The table above, with the results from the modified Wald test, strongly indicates that heteroskedasticity prevails. A high chi2 value[85] and a low P value[86] ("Prob>chi2") indicate that there is heteroskedasticity. In this case the chi2 value is extremely high and the P value is extremely low so the conclusion can be drawn with really high certainty. The regression must be adjusted for heteroskedasticity. Stata included all 290 municipalities in the test, so the test is reliable.

After it is shown that autocorrelation and heteroskedasticity prevails it is obvious to use a robust model, to deal with these problems, so a Fixed Effects model with robust standard errors is used.

8.3.3 Jointly significance

The regression is expected to produce jointly significant results. As a precaution, an F-test for jointly significance, produced together with the regression results in Stata, is taken into account.

Table 13: Jointly significance test results

F(16, 289)	12.208
prob>F	0.0508

[83] Shows the mean regression sum of squares divided by the mean error sum of squares.
[84] The probability of obtaining an F value as large as the one observed, given that all parameters except the intercept are zero.
[85] A measure of the goodness of fit.
[86] The probability of obtaining a chi2 value as large as the one observed, given that all parameters except the intercept are zero.

A high F value[87] and a low P value[88] ("Prob>F") indicate that the explanatory variables are jointly significant. The explanatory variables in the regression are jointly significant on a significance level of 5.08%. Together, the explanatory variables have an explanatory value. Stata included 289/290 municipalities and all explanatory variables in the test, so the result is reliable.[89]

8.4 Explanation of table

The effects of all the explanatory variables will be shown in tables with the headlines you can see below.

Coef.	Std.Err.	t	P>|t|	[95% Conf. Interval]

"Coef." Is the value of the variable's coefficient (Princeton University Library 2007). This is a measure of the estimated effect that the explanatory variable has on the explained variable. It should, however, be taken into account that the impact of the variable also is dependent on the variable itself. It does not matter if the absolute number of the coefficient is big, if the variable's difference between municipalities is small. It should also be mentioned that the absolute number is strongly affected by the unit used for the variable. If million persons and not thousand persons had been the unit for population, the coefficient would have been 1000 times as high. Given the characteristics of the variable, it is better the higher the coefficient is, since that makes it more likely that statistically significant conclusions can be drawn.

"Std. Err." is the standard error of the variable, how much the standard deviation varies across cases (Princeton University Library 2007). A low value of the standard error is desirable, since that makes it more likely that statistically significant conclusions can be drawn.

"t" is the t value, computed as follows (Princeton University Library 2007):
"t"="Coef."/"Std. Err."
The higher the absolute value of "t" is, the more likely that a statistically significant conclusion can be drawn.

"P>|t|" is the P value (Princeton University Library 2007). The P value is the measure of statistical significance, the probability of getting a result as extreme as the one obtained, in a collection of random data in which the variable has no effect.[90] The statistical significance is key in this study, since a low significance level is necessary to be able to draw any conclusions. The reasonable statistical level, that I will refer to later, means no more than 10% significance.

[87] Shows the mean regression sum of squares divided by the mean error sum of squares.
[88] The probability of obtaining an F value as large as the one observed, given that all parameters except the intercept are zero.
[89] The number is the same in the regression.
[90] "As extreme as" means as large "t" value as.

“95% Conf. Interval” means a 95% confidence interval. One can be 95% confident that the real, underlying value of the coefficient that is estimated falls somewhere in the 95% confidence interval (Princeton University Library 2007).