Effect of National Diversity on Football Team's Success

by

Polad Gulushov

Submitted to

Central European University

Department of Economics and Business

In partial fulfillment of the requirements for the degree of Master of Arts

Supervisor: Professor Laszlo Matyas

Budapest, Hungary

2020

I, the undersigned [Polad Gulushov], candidate for the degree of Master of Arts at the Central European University Economics and Business, declare herewith that the present thesis is exclusively my own work, based on my research and only such external information as properly credited in notes and bibliography. I declare that no unidentified and illegitimate use was made of work of others, and no part the thesis infringes on any person's or institution's copyright. I also declare that no part the thesis has been submitted in this form to any other institution of higher education for an academic degree.

Budapest, 16 June 2020

_____

Signature

Abstract

This thesis analyzes the relationship between national diversity and the achievements of football teams. Based on data from transfermarkt.com, consisting of players from 124 teams competing in the English, French, Spanish, Italian and German leagues for the seasons from 2016/17 to 2018/19, it shows that cultural diversity has an significant positive effect on a team's success, however, the effect is very small. This result can be used as a guide for the team management on signing the new players.

Table of Contents

Chapter 1

Introduction

Until 1995, to receive a Ballon d'Or, the most prestigious award that a footballer can ever earn, the player had to play for a European Club and have European citizenship. As a result, some of the best players of the 1950s to the 80s, such as Diego Armando Maradona, Pele never received this award (FIFA.com, 2014). Ever since Ballon d'Or has been won by non-Europeans several times. In fact, Argentinian Lionel Messi is known as the player who won this award the most. This change and the other significant event called the Bosman ruling in 1995 simulated the flow of players from different parts of the world to Europe.

Nowadays, most football teams contain players from all over the world. To see this, consider the Champions League Final match in 2018 between Real Madrid representing Spain and Liverpool representing England. Despite being two of the leading countries in football, only a few of the players on both sides were from the represented countries. In the 18-player squad of Real Madrid there were only seven Spanish players, while six for played Liverpool. In total, there were footballers from 15 countries, such as Costa Rica, Spain, France, Brazil, Croatia, Germany, Portugal, Wales, The Netherlands, England, Scotland, Egypt, Senegal, Belgium and Estonia. Apart from players the teams were also managed

by French and German coaches (UEFA.com, 2018).

These days, not only in football, but also in business areas, managers are encouraged to hire people from diverse nations and races for growth. (Llopis, 2012). Considering these well-known facts, the effect of national and cultural diversity on team success has been a hot topic for researchers to analyze. There have been both quantitative and qualitative research papers on the effect of diversity in a team on team's performance, namely, its productivity and success. The authors have used different datasets and research methods to analyze this issue, and they came to various conclusions. Some found positive relationship between the diversity in a working team and the productivity of that team (Ayega and Muathe, 2018; Ashikali and Groeneveld, 2015). This positive effect comes from the fact that different ethnic groups, nations have different approach to the problem, which increases the creativity in a workplace (Martin, 2014). Moreover, Dastane and Eshegbe (2015) showed that gender and ethnic differences have positive effect on employee satisfaction in a workplace. One notable thing in the literature (McMahon, 2010) is that the relationship between cultural diversity and the team success is U-shaped, which means team is successful when there is either high or low degrees of diversity.

My aim in this thesis is to analyze the impact of the diverse team players on the success of the team. The results of this analysis may be a guide for the football managers on signing new players and increase the efficiency of the game play. The rest of the thesis is organized as follows. In the next section I reviewed already existing research on this topic. In chapter 3, I gave the description of the data and the methodology that I used in my analysis. In chapter 4, I talked about the results of the analysis. The last chapter gives brief information about the limitations of the research, policy recommendations and possible improvement of the analysis.

Chapter 2

Literature Review

In this section, I summarize the key papers that helped me to develop my hypothesis and methodology. First, I talk about the papers, which found the positive effect of the team diversity. In contrast, in the second part, I mention the paper that showed the negative effect of the heterogeneity in the team. Finally, the papers I note in the last section found no significant relationship between success of the team and the diversity in the team.

2.1 Cultural Heterogeneity in the Team is Positively Related with the Team Success

Acheampong et al. (2019) found that, in the Africa Cup of Nations (AFCON), countries with the players from different football leagues perform the best. They used the team level data from AFCON participants between years 1990 and 2017, which is, in total, 213 observations. As a dependent variable, occasions when the medals were won was chosen. Using the probit model the authors concluded that the team consists of player who plays

at different leagues than local league does better than the team which does not. As a variable of interest, Acheampong et al. (2019) chose the ratio of foreign to local league players. Moreover, they also showed that having the players from different age groups also increases the probability of participation in the cup successfully.

Unlike the previous paper, they used different dataset and methodology. Based on Champions League (CL) data from 2003 to 2013, the authors chose goal difference as a proxy for the team success. For the variable of cultural diversity, they chose the Automated Similarity Judgment Program (ASJP) index. The ASJP index calculates the similarity between the languages based on the most used words in the both languages and produces the number between 0 and 100 as an index, 0 being the closest and 100 being the most distant. Using Structural Equation Model and implying Instrumental Variable estimation, they found that one-standard deviation rise in cultural heterogeneity rises per-game goal difference by 0.33 goals. It is also worth to mention that as an Instrumental Variable, they used the number of players from the same country.

Similar to football, the effect of diversity on team success has been studied in other sports too. Based on the data collected via questionnaires from 140 National Collegiate Athletic Association basketball coaches, Bell and Riol (2017) claim that the effect of cultural heterogeneity in a team is multidimensional and cannot be summarized with just one variable. The analysis found that cross-cultural communication competence has positive effect on collective efficiency and team effectiveness, while the other dimensions of the cultural differences do not significantly affect the team performance.

In parallel to football and basketball, similar research has been conducted for hockey too. Based on data from National Hockey League (NHL) from season 2001-2002 to season 2007-2008 excluding 2004-2005 season because of the rule changes specific to that particular season, citetkahane2013effects found that having a European hockey player in the NHL increases the chances of winning a match. For the dependent variable they chose the

percentage of the matches won and the percentage of the possible gainable points claimed. The diversity of the team was depicted in the model in the following way. The players were categorized in five groups: North America (players from U.S. and Canada), Czechoslovakia (players from either Czech Republic or Slovakia), Sweden, Finland, and Russia. The authors have calculated Hernahl-Hirschman index conditional on number of the Europeans in the team to control vast majority of players from the first category. Using the Fixed Effect regression, Bell and Riol concluded that there is a significant positive relationship between team success and the diversity in a team. However, the noteworthy point of this finding is that the Europeans in the team should be from the same category, otherwise, the team does not do well in the NHL.

2.2 Cultural Heterogeity in the Team is Negatively Related with the Team Success

Based on the data from top five football leagues (England, France, Germany, Italy, Spain) for the season 2008/09, which consists of 2483 player and 98 clubs, Maderer et al. (2014) found that there is a negative effect of cultural heterogeneity. To measure the success of the team, they chose a relatively different variable than the ones that I mentioned previously, which is 'Points to Market Value ratio' (PMVR). Based on previous literature (Brandes et al., 2009; Hall et al., 2002), the authors agree that the financial power of the club is related to the success of the club. Thus, they chose to use such variable as a success measurement. For the variable of cultural diversity, they used defined cultural heterogeneity in the following way:

$$CulturalHeterogeneity = \sum_{i=1}^{n}(P_i * ln(P_i)) * (-1) \tag{2.1}$$

In this equation, P_i denotes the proportion of the team that belongs to the nationality i. At the end of the regression analysis, not only the negative effect of the cultural heterogeneity, but also negative effect of the international coach was found.

Similarly, Chatman and Flynn (2001) found that the effect of demographic heterogeneity has negative effect on cooperation in the working teams. However, they also found that this negative effect disappears through time. The authors suggested to imply some norms to form the co-operations in the work places, based on Social Categorization Theory. Data was collected from 119 MBA students from different American universities. for the cultural heterogeneity, they chose the Euclidian Distance method.

Riordan and Shore (1997) also used the Euclidean Distance method to analyze the relationship between demographic heterogeneity and the attitudes of employees in the work units. Their work is based on survey data collected from 1584 insurance company employees in the U.S. The research yielded that people work with people from the same ethnicity or the same race better. As a dependent variable, the authors used different variables, namely, work group productivity, cohesiveness, commitment, and advancement opportunities. Proving McMahon's (2010) point that the relationship between cultural heterogeneity and the effective work is not linear, they showed that in Hispanic groups 50/50 working environment yields the best results. The researchers also found that gender heterogeneity has no significant impact on the effective work environment.

Haas and Nüesch (2012), similar to the other researchers, found that there is a negative effect of diversity on the team performance of Bundesliga teams. They reached this conclusion based on German league data from 1999/2000 to 2005/2006 season. The researchers implied three models, for the three different dependent variables, namely, points of the teams received, goal difference, which is the difference between goals the team scored and conceded, and the expert evaluation, which is the performance evaluation of the players, who played more than 30 minutes in the game using German grading scale, 1.0 being very

good, 6.0 being terrible. For the diversity variables, different from the other papers, they chose Blau index (Blau, 1977), calculated in the following way:

$$CulturalHeterogeneity = 1 - \sum_{i=1}^{n} p_i^2 \tag{2.2}$$

where p_i denotes the proportion of the players from country i. Finding the negative relationship between national heterogeneity and the team success, the authors claim that the language barrier is the biggest challenge that negatively effect the team play.

2.3 Cultural Heterogeneity Has no Significant Impact on the Team Success

Using the model proposed by Lazear (1999), Brandes et al. (2009) found that national diversity and team production are not significantly related. They used data from Bundesliga for the years from 2001/2002 to 2005/2006, in total 90 observations. As a dependent variable, the rank of the team at the end of the season was chosen. The authors only included the footballers, who played at least half of the season to their sample. For the variable of heterogeneity, they chose the number of nationalities in the core part of the team. Apart from national diversity, Brandes et al. also analyzed the country-specific skills. For instance, the empirical analysis showed that having a Brazilian player in midfield or in the attacking position boosts team's chances of scoring goals. They also found that having a multiple nationality in the defense position lets team concede more goals.

After analyzing more than hundred lab and field studies between 1959 and 2016, Guver and Motschnig (2017) claimed that the diversity does not necessarily effect the team productivity and success. They noted that the diversity in the workplace can be interpreted in two ways. Diverse workplace can create problems in communication, integration and

turnover, but it also creates an environment for creative and diverse approaches to the problem solving.

Another research that chose Blau's heterogeneity index was conducted by Omankhanlen and Ogaga-Oghene (2011). Based on data collected from 120 working teams across 62 bank branches in Lagos Island and Mainland Nigeria, the authors found that the ethnic diversity negatively, but insignificantly effects the organizational effectiveness. Moreover, tenure has positive insignificant effect on organization's effectiveness. In addition to those, gender diversity has negative, age diversity has positive significant effects on productivity of the working team. They chose productivity per employee to measure the performance criterion.

Like the papers I mentioned previously, Nüesch (2009) also used the Bundesliga data for six seasons from 2001/2002 to 2006/2007. He analyzed the effect of age, race and tenure heterogeneity on team performance under different models. The author chose points collected and goal differences for the team performance indicators. To measure the ethnic variety, Nuesch also chose Blau's index, and found that demographic effects are spurious. Holding the means of demographic indicators, none of the covariates has significant effect of team performance.

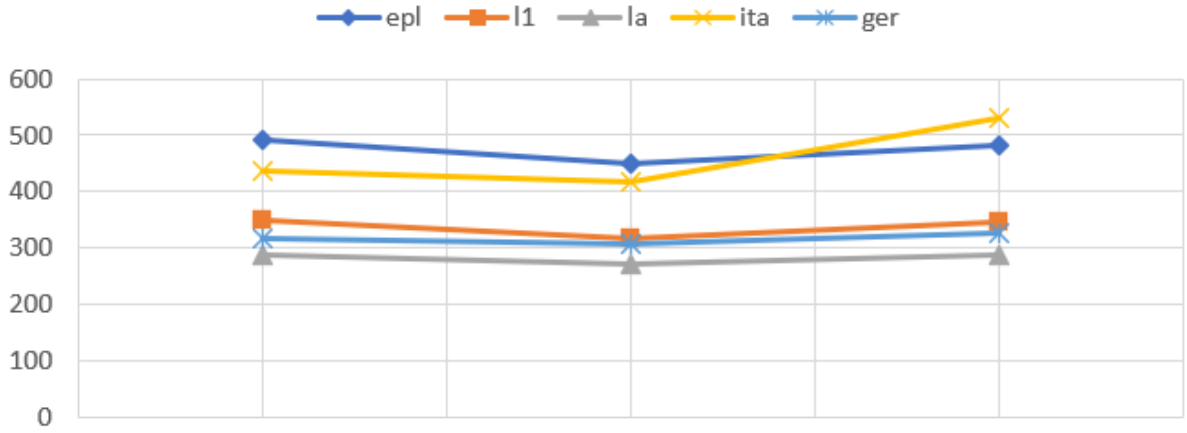
Chapter 3

Data and Methodology

My data is retrieved from trnasfermarkt.com and whoscored.com, which are some of the most reliable web sites for both individual and team level data on football. The dataset consists of details of the Big Five leagues: English Premier League (EPL), French League (Ligue 1), Spanish league (La Liga), Italian League (Serie A) and German Bundesliga (Bundesliga) from the last three completed football seasons from 2016/17 to 2018/19. In total, I used the data of players from 124 teams from these leagues. I particularly chose these leagues because they generate more than half of the total sales in whole European football (Deloitte, 2011) and had the most spectators in the whole world (Khanna, 2018).

Before I present the summary statistics of data, I would like to state some interesting facts about it. In five leagues, there are representatives from 120 nationalities. In 2016/17 season, Spanish players were dominating the Big Five leagues with over 500 players in total this. By the end of my sample timeline, number of Italian players increased rapidly. In 2018/19 season there were 666 Italians competing in these leagues. However, most of these players were competing in home country, while Spanish players are always part of the core team, who play more than 10 games per season. Another noteworthy thing is that

Figure 3.1: Dynamics of Foreign Players in Big Five Leagues



all these leagues are very "conservative", in a sense that there are very few players from other members of the Big Five. For example, in the French league only 2 of 754 footballers playing are from England in the last season of my sample data, while there were none in 2016/17. Looking at the raw data, one also can see that the French league has the highest proportion of the African players, which is not surprisingly, as we can see most of the French national team members are immigrants from African countries. Checking the other leagues, we can also observe that Serie A has the most portion of players from the Balkans and Latin America, while the Bundesliga has the most non-German population from Central Europe. La Liga has the teams with least diversity. In fact, the team with the least diversity index, Atletic Bilbao, competes in the Spanish league. In the Premier League, however, I did not observe any preference towards some regions or country. In above graph, the dynamics of foreign players in each league can be seen. One can see that Spanish La Liga always has the lowest number of foreigners, while English Premier League and Italian Serie A have changed their places over the course of time.

Descriptive Statistics

Variable	N	Mean	SD	Min	Max
Average age	98	24.108	1.058	21.6	27.1
Foreign Coach	98	.306	.463	0	1
Market Value	98	261.999	273.57	37.35	1106
Foreign Experience of the Coach	98	.296	.459	0	1
Age of Coach	98	49.418	7.658	30	70
Number of red cards	98	2.592	1.752	0	9
Goals	98	50.245	15.731	10	95
Number of players	98	41.408	10.672	29	88
Diversity	98	.695	.14	.112	.922

Table 3.1: Descriptive Statistics

Having known these facts now I can talk about the description of data, the variables I chose and their properties. In the following table, we see the descriptive statistics of the variables I chose. I did not include country specific variables, as it would be a table with more than 120 rows.

Next, I would like to discuss about the independent variables of the model.

- **Average age** is closely related to the team success. The younger the team members are, more stamina they have. However having all the players very young in the team is also not a good idea, as experience of the player is also key factor in winning.

- **Foreign coach** is a Dummy variable that takes value 1 if the team coach has different nationality than the the league the team is competing. Following Maderer et al. (2014), I included variable, because having a foreign coach show the flexibility of the team in terms of working with other nations.

- **Foreign Experience of the Coach** is also important factor, because it shows

coaches ability to work with multinational team. Similar to **Foreign Coach**, it is also a dummy variable that takes value one if the coach of the team has worked in at least two different leagues.Age of

- **Age of Coach** is also an important factor, as age includes both coaches experience and the communication with the other generation, which is closely related to the strong teamplay. Coach at different ages can have different successes, for instance, Mustafa Denizli, who was the manager of the Galatasaray from Turkey in 1989, took the team to the Champions League Semi Final. This is still the biggest success that any Turkish team has ever achieved in the respective tournament. However, in 2017, Galatasaray hired him, but was fired several months after the signing as the team was doing worse and worse every week.

- **Number of red cards** is the variable that meause number of red cards received the team during the season. Obviously it directly affects the team, because it creates disadvantage against winning the game.

- **Number of Players** is also important factor, as the bigger the team is, the more probability that there are more quality players. Have many players also gives an advantage of replacing injured players very easily.

- **Diversity** is measured by Blau index, given by equation (2.2). The closer to 1 index is, the more diverse the team is.

I also included some variables that are no on the list like market value of the team, league dummies as an independent variables.

For the dependent variables, in other words to measure success, I chose three variables.

- **Points** is the main measure of the success in the national leagues, the champion is the team with the most points at the end of the season.

- **Goals** measures the number of goals team scored during the season. I preferred this measure rather than goal difference is that goal difference can be misleading and I will show two examples from real life to explain this statement. The first example is from La Liga, where despite being the second most points, Atletico Madrid have scored less goals than some of the teams scored less points than them. The second example is from Latin America. In 1985 Coritiba became champion with the negative goal difference, scoring only 25 goals in 29 games (Langley, 2017). Therefore, I chose number of goals scored as a measure of good team play.

- **Euro** is a dummy variable that is equal to one if team successfully manages to get the right to participate in European cups.

Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
(1) Number of Players	1.000								
(2) Average Age	-0.389	1.000							
(3) Market Value	0.161	-0.130	1.000						
(4) Goals	0.057	-0.248	0.629	1.000					
(5) Foreign Coach	-0.069	-0.045	0.292	0.160	1.000				
(6) Foreign Experience of a Coach	-0.024	-0.039	0.219	0.142	0.568	1.000			
(7) Age of a Coach	0.027	0.058	0.000	0.013	-0.012	0.038	1.000		
(8) Number of red cards	0.191	-0.157	-0.211	-0.222	-0.107	-0.065	0.051	1.000	
(9) Diversity	-0.016	-0.015	0.016	0.069	0.061	0.058	0.012	0.025	1.000

Table 3.2: Correlation Matrix

Looking at Table 3.2, one can see that there is no significant correlation between the variables that can cause multicollinearity. Having talked about the variables, it is time to present the methodology that I used in the analysis. I conducted analysis in two parts. In the first part, I used cross-sectional data from the last completed season 2018/19. I implied multivariable linear regression for dependent variables goals, points and rank. For

the variable euro I used two different models, namely, Linear Probability Model (LPM) and Probit regression as it is binary variable. In the second part I combined all three seasons to conduct panel data analysis of the topic. Firstly, I used pooled OLS regression to find the effect of the national diversity on success. Later, I tried to find team fixed and random effects.

Chapter 4

Analysis and Discussion

As I mentioned in the previous chapter, I used several variables as a dependent variable and applied different models. This chapter consists of several sections each of which discuss different models.

4.1 Cross-sectional Analysis

4.1.1 Goals as a dependent variable

	(1) goals	(2) goals	(3) goals	(4) goals
Number of players	0.0881	0.100	0.0758	0.0756
	(0.58)	(0.64)	(0.48)	(0.48)
Average Age	-2.355*	-2.231	-2.110	-2.110
	(-2.10)	(-1.96)	(-1.81)	(-1.80)
Market Value	0.0304**	0.0293**	0.0304**	0.0304**
	(4.48)	(4.14)	(4.36)	(4.52)
Age of a Coach	0.261	0.262	0.244	0.244
	(1.65)	(1.63)	(1.52)	(1.51)
Number of red cards	-2.425**	-2.487**	-2.399**	-2.397**
	(-2.73)	(-2.72)	(-2.71)	(-2.72)
Diversity		6.618	9.793	9.891
		(0.83)	(1.13)	(1.05)
Foreign Coach			-3.149	-3.090
			(-1.07)	(-0.97)
Foreign experience of a coach				-0.140
				(-0.04)
_cons	88.76**	81.12*	78.35*	78.33*
	(2.80)	(2.39)	(2.27)	(2.26)
N	98	98	98	98

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$

Table 4.1: Regression Results: Goals

In table 4.1 I present the results of the regressions, where I chose the number of goals scored during season as a dependent variable. For this case, I include 4 different regression models. In column 1, the model with no heterogeneity variable is presented. As we can see the market value of the team and the number of red cards received have significant effects on the number of goals conceded. Going to the columns on the left I add diversity variables one by one. As one can see the positive coefficient of Blau's index (diversity) on number of goals scored increases as we add more variables, however, we never have statistically significant results. Interestingly, the effect of the foreign coach has negative effect, as we add more variables, similar to the Blau's index, it is still insignificant. Finally, the foreign experience of the coach has a positive, yet small (having a coach who worked abroad has less than one goal effect) and insignificant effect on number of goals.

4.1.2 Points as a dependent variable

I add the results of the regression using points scored during the season as a dependent variable to Table 4.2.

	(1) points	(2) points	(3) Points	(4) points
Number of Players	-0.0737	-0.0881	-0.0968	-0.0898
	(-0.55)	(-0.67)	(-0.72)	(-0.67)
Average Age	-3.159**	-3.307**	-3.264**	-3.248**
	(-2.87)	(-3.09)	(-3.03)	(-3.01)
Market Value	0.0451**	0.0465**	0.0468**	0.0460**
	(9.58)	(8.95)	(8.49)	(8.38)
Age of Coach	0.351**	0.351**	0.344**	0.348**
	(2.83)	(2.88)	(2.78)	(2.80)
Number of Red Cards	-0.962	-0.887	-0.856	-0.900
	(-1.46)	(-1.36)	(-1.32)	(-1.45)
Diversity		-7.936	-6.802	-9.626
		(-1.03)	(-0.85)	(-1.13)
Foreign Coach			-1.124	-2.807
			(-0.40)	(-0.86)
Foreign Experience of a Coach				4.007
				(1.35)
_cons	103.6**	112.8**	111.8**	112.5**
	(3.47)	(3.90)	(3.85)	(3.89)
N	98	98	98	98

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$

Table 4.2: Regression Results: points

One can observe similar results: market value and the number of red cards still have significant results. However, unlike the previous model we have more significant results. The structure of the table is the same as Table 4.1. We can see that apart from market value and the number of red cards, the coefficient of age of coach is significant. As I

add more diversity variables, I do not lose any of the significant variables. Moreover, the intercepts are also significant in our models in this section. Unfortunately, none of the diversity variables gain significance after switching to this method. Surprisingly, however, the coefficient of the diversity variable has different trend than the previous model. It increases after adding foreign coach variable, then decrease even further after adding the last variables. In the same model, having a foreign coach also has negative yet insignificant result. I would interpret it as the communication problems due to the language barrier.

4.1.3 Euro as a dependent variable

Linear Probability Model

	(1) euro	(2) euro	(3) euro	(4) euro
Number of Players	0.000843	0.000939	0.000266	0.000201
	(0.24)	(0.27)	(0.07)	(0.05)
Average Age	-0.123**	-0.122**	-0.119**	-0.119**
	(-3.19)	(-3.15)	(-3.04)	(-3.06)
Market Value	0.000939**	0.000930**	0.000959**	0.000966**
	(8.90)	(8.20)	(7.65)	(7.58)
Age of Coach	0.0105*	0.0105*	0.0100*	0.01000*
	(2.40)	(2.37)	(2.20)	(2.19)
Number of Red Cards	-0.0734**	-0.0739**	-0.0714**	-0.0710**
	(-3.72)	(-3.69)	(-3.47)	(-3.41)
Diversity		0.0528	0.141	0.168
		(0.23)	(0.61)	(0.66)
Foreign Coach			-0.0873	-0.0714
			(-1.10)	(-0.87)
Foreign Experience of a Coach				-0.0379
				(-0.38)
_cons	2.710**	2.649*	2.573*	2.565*
	(2.64)	(2.51)	(2.43)	(2.43)
N	98	98	98	98

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$

Table 4.3: Regression Results: lpm

In this subsection, I discuss the Linear Probability Model, using the right to participate in the European Cups as a dependent variable. In Figure 4.3, I present the regression results. In a similar fashion Market Value and the number of red cards received have significant effect of the success of the team. Moreover, the age of the coach also has significant positive yet small (less than 1 percent) effect on the success of the team. Similar to the previous models, we still cannot get statistically significant coefficients for the diversity variables. The Blau's index has positive effect on the probability of participating in the European Cups and it increases as more variables are added. Having a foreign coach and foreign experience of the coach still have a negative and insignificant effect on the dependent variable. Like previous model, we have significant coefficient for the intercept.

Probit

	(1) euro	(2) euro	(3) euro	(4) euro
Number of Players	-0.000458 (-0.02)	-0.0000937 (-0.00)	-0.00376 (-0.19)	-0.00370 (-0.19)
Average Age	-0.570* (-2.18)	-0.568* (-2.17)	-0.560* (-2.09)	-0.558* (-2.09)
Market Value	0.00542** (4.68)	0.00534** (4.65)	0.00579** (4.84)	0.00578** (4.90)
Age of Coach	0.0466* (2.40)	0.0465* (2.38)	0.0465* (2.33)	0.0462* (2.31)
Number of Red Cards	-0.527** (-2.71)	-0.531** (-2.72)	-0.536** (-2.66)	-0.538** (-2.66)
Diversity		0.318 (0.25)	1.216 (0.77)	1.199 (0.75)
Foreign coach			-0.720 (-1.65)	-0.771 (-1.56)
Foreign Experience of a Coach				0.0861 (0.19)
_cons	10.93 (1.69)	10.68 (1.67)	10.10 (1.57)	10.07 (1.57)
N	98	98	98	98

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$

Table 4.4: Regression Results: probit

The next model of this thesis is a generic probit model, using the same variable from the previous subsection as a dependent variable. Figure 4.4 show the marginal effects of

the variables on the participation in the European Cups. In terms of significance there is almost no difference from the model discussed above. The Market Value, the Age of Coach and the number of red cards still have their significant effect. Similarly, diversity variables still have insignificant effects on the probability of participating in the European Cups. Blau's index still has a positive increasing, yet insignificant effect on the dependent variable. The Foreign coach insignificantly decreases the chances of participating in the Eurocups.

4.2 Panel Data

In this subsection, I extend the analysis to the panel data methods. Similarly, I divide this subsection into even smaller ones dedicated to each dependent variable.

4.2.1 Goals as a dependent variable

To conduct Cross-sectional analysis, I started with the model with no diversity variable and added the variables one by one. For simplicity, I skipped those part and directly started with the model, where all variables are included. I, firstly, ran pooled OLS regression. The results are stored in the first column of the table, p-value significantly decreased from cross-sectional case, but the results are still statistically insignificant. Secondly, I used robust standard errors for regressions. Finally, I had significant coefficient of diversity, but the effect of the diversity is so small that it is negligible. After checking the successful teams, one can argue that it is reasonable to check for team effects. To choose between fixed and random effect models, I ran the Hausman test. Critical value is 156.56, we reject the null hypothesis that preferred model is fixed effect model. So, I choose fixed effect model. I lost the significance of the coefficient of interest.

	Pooled OLS	Robust Standard Errors Model	Fixed Effect Model	Clustered Standard Errors Model
Dependent Variable	goals	goals	goals	goals
Number of Players	-0.264*	-0.264*	-0.00503	-0.00503
	(-2.31)	(-2.18)	(-0.03)	(-0.02)
Average Age	-4.146**	-4.146**	0.232	0.232
	(-4.87)	(-4.99)	(0.15)	(0.17)
Market Value	0.0502**	0.0502**	-0.0426**	-0.0426**
	(12.54)	(7.53)	(-4.35)	(-2.70)
Foreign Coach	-2.345	-2.345	-1.410	-1.410
	(-1.14)	(-1.04)	(-0.50)	(-0.50)
Foreign Experience of a Coach	0.767	0.767	-0.186	-0.186
	(0.38)	(0.35)	(-0.07)	(-0.06)
Age of a Coach	0.0854	0.0854	0.201	0.201
	(0.76)	(0.83)	(1.31)	(1.58)
Number of Red Cards	-0.951*	-0.951**	-0.879	-0.879
	(-2.53)	(-2.60)	(-1.82)	(-1.92)
diversity	0.00814	0.00814*	0.00508	0.00508**
	(1.36)	(2.57)	(0.86)	(2.72)
_cons	151.0	151.0**	48.15	48.15
	(6.57)	(6.65)	(1.12)	(1.18)
N	292	292	292	292

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$,

Table 4.5: Goals as a Dependent Variable

For the last model, I used clustered standard errors. The effect of the diversity decreases even further but now it is significant at five percent level.

4.2.2 Points as a dependent variable

	Pooled OLS	Robust Standard Errors Model	Fixed Effect Model	Clustered Standard Errors Model
Dependent Variable:	points	points	points	points
Number of Players	-0.317**	-0.317**	-0.268*	-0.268
	(-3.05)	(-2.95)	(-2.07)	(-1.53)
Average Age	-3.821**	-3.821**	0.357	0.357
	(-4.96)	(-4.86)	(0.29)	(0.28)
Market Value	0.0561**	0.0561**	-0.0122	-0.0122
	(15.47)	(11.75)	(-1.61)	(-1.20)
Foreign Coach	-2.555	-2.555	-3.031	-3.031
	(-1.37)	(-1.21)	(-1.40)	(-1.32)
Foreign Experience of a Coach	2.633	2.633	2.851	2.851
	(1.43)	(1.32)	(1.42)	(1.16)
Age of a Coach	0.139	0.139	0.184	0.184
	(1.37)	(1.48)	(1.56)	(1.93)
Number of Red Cards	-0.642	-0.642*	-0.897*	-0.897*
	(-1.89)	(-2.04)	(-2.40)	(-2.54)
Diversity	0.00624	0.00624*	0.00256	0.00256**
	(1.15)	(2.04)	(0.56)	(3.25)
_cons	139.9**	139.9**	49.58	49.58
	(6.72)	(6.73)	(1.48)	(1.47)
N	292	292	292	292

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$

Table 4.6: Points as a Dependent Variable

I conducted similar analysis, for the points as well. I conducted Hausman test to choose between fixed and random effects. Critical value is 213.92, therefore, we will choose fixed effect model. The results of the models I used for points is indicated in the Table 4.7. Similar to the case with goals the effect is so small, even though it is statistically significant.

4.2.3 Time fixed effects

Out of curiosity, I also analyzed time effect seasonal effects. I ran Hausman test to choose to choose the fixed effect model, as critical value is 83.51.

	Time Fixed Effects	**Time Fixed Effects**
Dependent Variable	goals	points
Number of Players	-0.105 (-0.43)	-0.311 (-1.61)
Average Age	0.0711 (0.05)	0.133 (0.11)
Market Value	-0.0437* (-2.34)	-0.00393 (-0.34)
Foreign Coach	-1.385 (-0.49)	-2.977 (-1.25)
Foreign Experience of a Coach	-0.246 (-0.08)	2.633 (1.08)
Age of a Coach	0.197 (1.56)	0.167 (1.68)
Number of Red Cards	-0.816 (-1.68)	-1.031** (-2.88)
Diversity	0.00552** (2.89)	0.00436** (3.68)
2016.season	0 (.)	0 (.)
2017.season	-2.623 (-1.59)	-2.906* (-2.38)
2018.season	0.0580 (0.03)	-2.824 (-1.60)
_cons	57.01 (1.28)	58.26 (1.71)
N	292	292

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$

Table 4.7: Time Fixed Effects

In the table above, I presented the results of the estimation. I tested to see if the

coefficients of season dummies are jointly equal to zero. Critical values for the models are 0.1523 and 0.0528 respectively, meaning we cannot reject the null hypothesis that the coefficients of seasonal dummies equal to zero. The coefficients of the diversity are statistically significant in both models, however, both numbers are quite small.

I will skip the part with LPM and Probit Models for panel data, as it need some more detailed analysis. From the models that I have conducted in this thesis, we observed that data of five leagues for one season does not provide enough evidence for the effect of the diversity on success. However, large scale analysis of the topic gives us statistically significant results. Unfortunately, the coefficients are so low that we can neglect them.

Chapter 5

Conclusion

In this thesis, I analyzed the effect of the national heterogeneity on the success of the team. To conduct the analysis I applied econometric methods, namely, simple linear regression, Probit regression for cross-sectional analysis. I also conducted panel data research on particular topic. The result showed that despite getting the positive coefficients for the models, there is no significant effect of the cultural and national diversity on the team's achievements on cross-sectional level. I also included more seasons into my sample to conduct fixed effect estimates and get more precise estimates. I got the significant positive effect of the national heterogeneity on a team success. Problem with the results is that they are significantly small, so that we can almost neglect them. First limitation of the research was the size of the sample that was analyzed. I chose to analyze the big five leagues, because these leagues are on the center of the football arena. They have the most spectators and generate most revenues. Considering these facts usually footballers from the rest of Europe and from other parts of the world has interest in playing in these leagues, which decreases the level of national heterogeneity in the teams from other leagues. However, some teams are basically, too good, namely, if we check the champions from last few years in big five

leagues, we can see that champions usually do not change. For example, Bayern Munich have been winning Bundesliga for the last seven years, Paris Saint Germain has won six out of last seven seasons, and Italy have not seen a different champion than Juventus since 2011. The future research on topic can be extended by adding more leagues, where there are variations on the list of champions after careful considerations. Second and more "lethal" limitation of the thesis was the time constraint. Due to the reason I mentioned above, it would be more reasonable to make this research based on data for more than one year and calculate the team fixed effects to capture Since there is no data about the Blau's index of the football teams I had to calculate it by myself. However, doing so took so much time that I had to change the thesis content a lot. I could only analyze the topic for the last three seasons, but ideally, I would suggest doing the analysis for at least five year for more believable and important results. Another possible extension of the research would be looking at the individual level data and have more detail discussion about which countries have a stronger effect on the team play or which position suits the given country nationals the best. For example, Brazil has given the best strikers to the world of football, while Italians are famous for their systematic defense strategies and Spain has some of the best midfielders of all time. Such research would help the managers to address correct nationalities when it comes to signing new player.

References

Acheampong, E. Y., Akwaa-Sekyi, E. K., and Bouhaouala, M. (2019). How does team composition affect performance in continental tournaments? *Cogent Social Sciences*, 5(1):1606133.

Ashikali, T. and Groeneveld, S. (2015). Diversity management for all? an empirical analysis of diversity management outcomes across groups. *Personnel Review*.

Ayega, E. N. and Muathe, S. (2018). Critical review of literature on cultural diversity in the work place and organizational performance: A research agenda. *Journal of Human Resource Management*, 6(1):9–17.

Bell, R. and Riol, C. F. (2017). The impact of cross-cultural communication on collective efficacy in ncaa basketball teams. *International Journal of Cross Cultural Management*, 17(2):175–195.

Blau, P. M. (1977). *Inequality and heterogeneity: A primitive theory of social structure*, volume 7. Free Press New York.

Brandes, L., Franck, E., and Theiler, P. (2009). The effect from national diversity on team production—empirical evidence from the sports industry. *Schmalenbach Business Review*, 61(2):225–246.

Chatman, J. and Flynn, F. (2001). The influence of demographic heterogeneity on the emergence and consequences of cooperative norms in work teams. *Academy of Management Journal*, 44:956–974.

Dastane, O. and Eshegbe, J. W. (2015). Effect of diversity elements at workplace: An empirical study. *International Journal of Accounting and Business Management*, 3(1):62–69.

FIFA.com (2014). Welcome to fifa.com news - 14 facts about the fifa ballon d'or.

Guver, S. and Motschnig, R. (2017). Effects of diversity in teams and workgroups: A qualitative systematic review. *International Journal of Business, Humanities and Technology*, 7(2):1–29.

Haas, H. and Nüesch, S. (2012). Are multinational teams more successful? *The International Journal of Human Resource Management*, 23(15):3105–3113.

Hall, S., Szymanski, S., and Zimbalist, A. S. (2002). Testing causality between team performance and payroll: the cases of major league baseball and english soccer. *Journal of Sports Economics*, 3(2):149–168.

Khanna, A. (2018). Top 10: Most watched football leagues in the world, bundesliga tops the list.

Langley, D. (2017). Have any football teams won the league with a negative goal difference?

Lazear, E. P. (1999). Globalisation and the market for team-mates. *The Economic Journal*, 109(454):15–40.

Llopis, G. (2012). Diversity management is the key to growth: Make it authentic.

Maderer, D., Holtbrügge, D., and Schuster, T. (2014). Professional football squads as multicultural teams: Cultural diversity, intercultural experience, and team performance. *International Journal of Cross Cultural Management*, 14:215–238.

Martin, G. C. (2014). The effects of cultural diversity in the workplace. *Journal of Diversity Management (JDM)*, 9(2):89–92.

McMahon, A. M. (2010). Does workplace diversity matter? a survey of empirical studies on diversity and firm performance, 2000-09. *Journal of Diversity Management (JDM)*, 5(2).

Nüesch, S. (2009). Are demographic diversity effects spurious? *Economic Analysis & Policy*, 39(3):379–388.

Omankhanlen, A. E. and Ogaga-Oghene, J. (2011). The impact of workforce diversity on organizational effectiveness: a study of a nigerian bank. *Annals of the University of Petrosani, Economics*, 11(3):93–110.

Riordan, C. and Shore, L. (1997). Demographic diversity and employee attitudes: An empirical examination of relational demography within work units. *Journal of Applied Psychology*, 82:342–358.

UEFA.com (2018). All the 2017/18 champions league results.