

---

## Research-Thesis, COLLÉGIALITÉ

**Auteur** : Piccoli, Anthony

**Promoteur(s)** : Gautier, Axel

**Faculté** : HEC-Ecole de gestion de l'Université de Liège

**Diplôme** : Master en sciences économiques, orientation générale, à finalité spécialisée en economic, analysis and policy

**Année académique** : 2019-2020

**URI/URL** : <http://hdl.handle.net/2268.2/8834>

---

*Avertissement à l'attention des usagers :*

*Tous les documents placés en accès ouvert sur le site le site MatheO sont protégés par le droit d'auteur. Conformément aux principes énoncés par la "Budapest Open Access Initiative"(BOAI, 2002), l'utilisateur du site peut lire, télécharger, copier, transmettre, imprimer, chercher ou faire un lien vers le texte intégral de ces documents, les disséquer pour les indexer, s'en servir de données pour un logiciel, ou s'en servir à toute autre fin légale (ou prévue par la réglementation relative au droit d'auteur). Toute utilisation du document à des fins commerciales est strictement interdite.*

*Par ailleurs, l'utilisateur s'engage à respecter les droits moraux de l'auteur, principalement le droit à l'intégrité de l'oeuvre et le droit de paternité et ce dans toute utilisation que l'utilisateur entreprend. Ainsi, à titre d'exemple, lorsqu'il reproduira un document par extrait ou dans son intégralité, l'utilisateur citera de manière complète les sources telles que mentionnées ci-dessus. Toute utilisation non explicitement autorisée ci-avant (telle que par exemple, la modification du document ou son résumé) nécessite l'autorisation préalable et expresse des auteurs ou de leurs ayants droit.*

---

**Evaluation of determinants of performance of  
the clubs participating in the Belgian football  
championship**

Jury

A thesis submitted by  
**ANTHONY PICCOLI**

Supervisor:  
Axel Gautier

For the degree of Master in Economics,  
specialized in Economics Analysis and Policy

Readers:  
Jerome Schoenmaeckers  
Bernabe Walheer

Academic year 2019-2020

## Acknowledgments

I would like to thank the people who took part in the realization of this thesis. Without them, this project would not have been the same.

First and foremost, I am grateful to my thesis supervisor, Professor Axel Gautier. He gave me the opportunity to find an interesting topic. He provided me with databases, references, books, and tips. He organized a conference regarding this subject, and finally, he gave me a methodology in order to achieve this project.

I also wish to give my gratitude to my two thesis readers: Professor Jerome Schoenmaeckers and Professor Barnabe Walheer, for their advice, suggestions during the course "Master thesis Methodology", and their availability to hold appointments for me. The latter for the project during the course "Applied Microeconometrics" which was a real added value for having a clear overview of my work during my thesis. I am also thankful to M. Walheer for solving all my econometrical issues.

Additionally, I would like to thank Michel Lecompte, head of the Sports department at *La RTBF*<sup>1</sup> and above all, a specialist in Belgian football for more than 40 years. He was always eager to share his knowledge, references, and people to contact to get the critical pieces of information. May you have a good retirement.

I also want to express my appreciation to my friends, Benjamin, Elsa, Miguelina, Tristian, my cousin Jonathan, and my brother, Adrien. You reviewed this thesis and gave me valuable remarks.

Finally, it is essential for me to thank my parents. You both are my examples. This project is the consecration of all these sacrifices you have made, and I dedicate it to you.

---

<sup>1</sup>Public-service broadcasting in the French community in Belgium.

## Executive Summary

The purpose of this research<sup>2</sup> is to find empirical evidence about variables that might have any influence on football performance, but especially for Belgium. It contributes to prior studies for other countries in Europe and North America.

Literature regarding the subject has begun to grow at the beginning of the 21st century with the rise of data accessibility. Researchers intended to test different team characteristics such as in coaches' influence on teams, revenue, age, or financial aspect like wages, revenue, transfers, and so on. The major part discovered that salaries are positively correlated with success.

In order to perform our investigations, a panel data of twenty-three teams over ten seasons, from 2009-2010 to 2018-2019, is used. The advantage of using longitudinal data is that they remove bias that can be caused by omitted or unobserved variables. Furthermore, from this sample, a sub-sample is created. It includes teams from the K11 only. Then, a third sample, independent from the first two, analyses triumph in play-offs 1.

The presence of fixed effects in our model fosters the utilization of the Within estimate. Nevertheless, comparison thanks to Pooled OLS will be exposed as well as System GMM which permits inertia from the past in our model and increase the reliability of the estimates.

The findings come up with three mains conclusions: larger wage expenditures, additional games during the season, and better attendance from supporters enhance football triumph in Belgium. The two other data set to refine the results.

Therefore, we recommend to anyone in charge of football administration to highly invest in team payroll instead of transfers or to generate good financial incomes if the club is considered as a "sub-top" club. Then, not to neglect the bonuses of playing several competitions except if the club succeeds in qualifying for play-offs 1. Finally, to maintain a good relationship with the fans.

---

<sup>2</sup>Word count: 14 257

# Abstract

In football, some teams win while others lose. However, what really matters in this sport? This thesis tries to determine which variables could have a potential influence on team performance for the top Belgian football championship known as the *Jupiler Pro League*. To achieve this, an empirical analysis is undertaken thanks to a panel data from the 2009-2010 season to the 2018-2019 season. The results show that higher salaries expenditures, extra competitive games in European or national cups, and better attendance improve seasonal achievement. Nevertheless, when the sample is reduced to only clubs out of the five best Belgian teams, it appears that only more substantial financial incomes are positively correlated to seasonal achievement. Lastly, the analysis of the play-offs 1 since their implementation, reveals the effect of additional games becomes negative and high ranking at the end of the regular competition impact significantly triumph in that second stage of the competition.

**Keywords:** Football, soccer, sports performance, Belgium, panel data.

## Nomenclature

The next list describes abbreviations that will be later used within the body of the document:

- CL** Champions League (or C1)
- DV** Dependent variable
- EL** Europa League (or C3)
- FE** Fixed effects
- G5** Anderlecht, Club Brugge, Genk, Gent, and Standard de Liège
- JPL** Jupiler Pro League
- K11** Clubs others than those belonging to the G5
- OLS** Ordinary Least Squares
- PO** Play-offs
- RE** Random effects
- UEFA** The Union of European Football Association

# Contents

<b>1</b>	<b>Introduction</b>	<b>8</b>
<b>2</b>	<b>Literature review</b>	<b>10</b>
<b>3</b>	<b>Data</b>	<b>12</b>
3.1	The Jupiler Pro League . . . . .	12
3.2	Data collection and description . . . . .	13
3.3	Descriptive statistics . . . . .	17
3.3.1	Summary statistics . . . . .	17
3.3.2	Team payroll and team performance . . . . .	17
<b>4</b>	<b>Empirical approach</b>	<b>19</b>
4.1	Model . . . . .	19
4.2	Endogeneity check . . . . .	22
<b>5</b>	<b>Results</b>	<b>23</b>
5.1	Results for the total sample . . . . .	24
5.1.1	League position as the dependent variable . . . . .	24
5.1.2	Point percentage as the dependent variable . . . . .	26
5.1.3	Comparison . . . . .	27
5.2	Results for K11 sample . . . . .	30
5.2.1	League position as the dependent variable . . . . .	30
5.2.2	Comparison with total sample . . . . .	32
5.3	Results for play-offs 1 sample . . . . .	32
5.3.1	League position as the dependent variable . . . . .	33
5.3.2	Comparison with the total sample . . . . .	33
<b>6</b>	<b>Policy recommendations</b>	<b>35</b>
<b>7</b>	<b>Conclusion</b>	<b>36</b>
	<b>Bibliography</b>	<b>39</b>
	<b>Appendix</b>	<b>43</b>
A.1	Clubs who took part in the <i>Jupiler Pro League</i> since 2009-2010 until 2018-2019 . . . . .	43
A.2	Clubs who played the play-offs 1 since their introduction until 2018-2019	43
A.3	Financial evolution of the <i>Jupiler Pro League</i> . . . . .	44
A.4	Correlation matrices complete sample . . . . .	45

A.5 Relation team payroll on team performance when using point percentage as DV . . . . .	46
A.6 Additional regressions . . . . .	47

## List of Figures

1	Relation team payroll on team performance (1). . . . .	18
2	Relation revenue and ranking. . . . .	19
3	Relation team payroll on team performance (2). . . . .	46

## List of Tables

1	Variables used in the model and sources . . . . .	14
2	Summary statistics . . . . .	18
3	Regression table for regular competition (league position as the DV)	25
4	Regression table for regular competition (point percentage as DV)	28
5	Regression table for K11 sample . . . . .	31
6	Regression table for play-offs 1 sample . . . . .	34
7	Average financial evolution by season (in millions of euros) . . . . .	44
8	Correlation matrix with league position . . . . .	45
9	Correlation matrix with point percentage . . . . .	45
10	Decomposition of the main regression by pooled OLS (1) . . . . .	47
11	Decomposition of the main regression by pooled OLS (2) . . . . .	48

# 1 Introduction

Football<sup>3</sup> is one of the most beloved and watched sport worldwide. In certain countries, this game is even considered as a religion. This sport has the potential to build relationships between people, communities, and countries together. When we take a look at the figures from the last *FIFA World Cup 2018*, 3.572 billion people in total watched the competition (FIFA, 2018). This means that approximately one in every two persons on this planet turn on their TV or their devices to watch the World cup.

Furthermore, football is a tremendous business. It generates billions of euros every year. In Europe alone, "top-division" clubs reported total combined revenues of 21 billion euros (UEFA, 2020). Within this income stream, we find tickets sold for every game, sponsorships, TV-broadcasting, and *UEFA* pricing. Regarding this last one, it is becoming more and more critical for teams to qualify for the continental competitions such as the *Champions League* (CL) or the *Europa League* (EL). In fact, CL and EL distributed respectively 1.976 billion euros (UEFA, 2019b) and 559.2 million euros (UEFA, 2019c) for the 2018-2019 season. For instance, Liverpool, the winner of the last C1, earned more than 110 million euros as a UEFA prize. More impressively, only for the English Premier League, TV-rights were sold for more than 3.6 billion euros for the 2019-2020 season (UEFA, 2020). These immense amounts tend to make club spending more, alongside higher wages or enormous transfers exceeding now hundreds of millions of euros.

In 1995, the *Bosman* ruling changed the face of football forever. It allowed for out-of-contract footballers to negotiate and to move to other clubs without paying transfer fees to the former club. This rule also erased the *UEFA* restriction of "3+2" in their competitions, in which only three foreigners were allowed in a squad + two if the foreigners came from youth academy (Frick & Simmons, 2014). It was the starting point for the deregulation of wages, transfer fees, nationalities in competitions, and more. Nowadays, it is very widespread to see millions of euros spent on new players, exorbitant salaries, or even starting eleven without any domestic players, similar to that of Inter Milan in 2010 when they won the *UEFA Champions League* final. More recently, in 2009, a new regulation made its apparition in the football landscape: the *Financial Fair-Play*. The main purpose of this new legislation introduced by the *UEFA* was to border transfer acquisitions and financial losses (UEFAa, 2019). It was one of the first restrictions implemented by football authorities. Now, clubs are no more able to spend countless money under pain of punishment.

In Belgium, competition is following the European upward trend, as exposed in the last European Club Footballing Landscape by the *UEFA* (2020), where clubs

---

<sup>3</sup>In the United States, they call football "soccer".

generate more and more money thanks to TV-allowance, gate incomes, sponsorship, etc. For instance, in March 2020, the *Jupiler Pro League* (JPL), the Belgian first division, stroke a big new deal for their TV-rights. With this new deal, the league is projected to earn more roughly 103 million euros annually over the next five seasons (Soccerex, 2020). This new contract should allow the distribution of more revenue to the clubs participating in the JPL and perhaps, improve their performances in the long run by attracting new players, through the building of new facilities and so forth. It is the confirmation that again, football generates a lot of money, and broadcasters are ready to pay considerable prices to have the exclusivity on certain leagues. They know that everyone who does not have the possibility to attend a game in a stadium will make everything possible to watch their favorite teams.

The Belgian championship is often considered as a launching pad to access in the future, "top-level" championships. Therefore, clubs in the JPL have understood the importance of good management. Indeed, they try to attract young talents or unknown foreign players, keeping them mostly between two-three years and after, trading them with massive fees in order to make added-values. This mechanism allows them to stay competitive against championships considered as "better". For example, in the current 2019-2020 season, 60% of the players are non-natives (Transfermarkt, 2020) and during the last summer, teams traded them for more than 221 million euros and bought for 134 million euros (Gautier, 2019). A positive difference than confirms the importance of transfers in the Belgian scheme.

Hence, with all these elements, a question arises: which variables (or inputs in this case) have a significant influence on team performance in football? Several studies (see section 2) had already attempted to answer the question in the past and have granted us insight into different competitions and even other sports. Since the beginning of the 21st century, the literature concerning has grown fast. Nevertheless, any research focused on Belgium. We propose to fill this gap by adding our contribution through this thesis, with particular attention given to the highest football level in Belgium: the *Jupiler Pro League*.

To achieve these investigations, a panel data of twenty-three teams who participated at least once in the first division, since the 2009-2010 to 2018-2019 seasons, was collected. Since previous empirical studies mainly employed one independent variable in their models, we will regress several independent variables on team success with the purpose of controlling for biasness and identify the most relevant. Three data sets will be employed; the first one on data for the regular season, the second one, data on regular season without the G5 (which is the five clubs with best performances in Belgium over the last five years) and the last one corresponding to the teams participating in the play-offs 1 (PO1) since their introduction. As a result, the approach studies linear models to analyze static and dynamic panel data.

This thesis is organized in the following manner. First, section 2 exposes the currently existing literature related to this subject. Section 3 offers some explanations about the soccer competition in Belgium, describes where the data comes from, whereas the last part is dedicated to descriptive statistics. The methods chosen to develop the model alongside empirical analyses are explored in section 4, while section 5 presents the results obtained with interpretations and robustness checks. Finally, sections 6 and 7 conclude this paper with some policy recommendations and a conclusion.

## 2 Literature review

Assessing team performance in sports is not unheard of, even if the interest in football is recent. Economists have always been interested in sport economics. In the beginning, first researchers were concerned with variables that might have an influence on team performance focused mainly on North American sports, such as baseball, basketball, and American football. This is probably because that soccer takes time to generate results from economic analysis, and, in that time, available evidence and data could be weak.

Scully (1974), Marburger (1994), Simmons & Forrest (2002), Stiroh (2002), and also Annala & Winfree (2011) all tried to evaluate the importance of salaries, coaches, wage disparity, stats performance and counting on team performance. Their principal findings were that wage disparity is negatively correlated to team achievement (Annala & Winfree (2011)). Scully (1974) showed that salaries have a positive impact on seasonal performance, while Stiroh (2002) confirms this theory, adding coaching as an important factor too.

The end of the 20th century and the beginning of the 21st century marked finally the arrival of analysis related to football with Szymanski & Smith (1997) and Simmons & Forrest (2002). They displayed two major results; first, greater payrolls are associated with greater success and second, the latter paper compared the influence of wages on seasonal realization in soccer and in baseball, in which they found the impact of higher salaries on performance as more important in soccer than in other North American sports. Divers explanations can explain this aspect: on the one hand, league authorities in the United States and Canada act as a "welfare state". That is to say, in order to maintain competition, policymakers installed payroll taxes, salary caps, draft<sup>4</sup> system in favor of teams considered as "lower" teams. Then, the free-agent system, which is stronger than in European football, triggers a monopsony for players and has pushed clubs' wages upward

---

<sup>4</sup>The draft system in baseball, basketball, American football, consists of selecting best new young players (generally after graduation) by the worst teams in the championship from the past season. Conditions of eligibility depend on the leagues (MLB & NBA & NFL, 2020).

without necessarily increasing performance. Moreover, those leagues do not include a relegation-promotion structure and qualifications for continental competitions. Consequently, teams do not have to struggle against fear of going back to the second division or missing European tournaments.

After that, the literature concerning soccer performance grew quickly with bigger data availability (company such as *OPTA* provides information to clubs thanks to their works with data), more transparency from the clubs (annuals account published every year), and more reliable resources on the Internet. Reviews about determinants of international victory appeared in Hoffmann et al. (2002), Houston & Wilson (2002), and Gásquez & Royuela (2016). These papers found that wealth, climatic conditions, demography, access to facilities, and football promotion across the country are positives variables on international success. On the contrary, healthcare or football history does not have a significant impact.

Then, investigations regarding clubs had risen too. Economists wanted to determine for example the effect of wage bill distribution on team accomplishment (Franck & Nüesch, 2011; Buccioli et al., 2014; Caruso et al., 2016; Yamamura, 2015), the effect of foreigners players on team performance (Bachan et al., 2014), (Royuela & Gásquez, 2019) and other independent variables like, manager contribution (Muehlheusser et al., 2018) or playing different competitions (Picazo-Tadeo & González-Gómez, 2010). The outcome of wage dispersion and change in coaches during the season, as well as playing different leagues, to influence success negatively. In regard to non-native players, it has no effect. The important is to choose the good one. Another last variable included in Tillou and Liarte analysis (2012), was the emphasis on the effect of age. In reality, they wanted to check the effect of experience on performance. They discovered that teams with only old players are not a proof of success. The best deal is to have a majority of young players framed with a minority of old players. This combination permits the group to have the experience of the older and the freshness of the younger.

However, the most meaningful evidence that researchers have found is the influence of salaries. Indeed, most of the papers undertook studies to know if it is worth to spend billions of dollars on new incoming players instead of rather spending more on team payroll. In theory, the correct answer is the second one. In their book, Kuper & Szymanski (2018) confirm this theory by analyzing the championship in England. As noted: "The higher the total wage bill, the better the squad, the higher the team will finish in the league. [...] Soccer clubs need to make fewer transfers.". Obviously, they are not alone in arguing this. As mentioned earlier, Szymanski & Smith (1997) and Simmons & Forrest (2002) were the first to analyze football and discovered the same conclusions for the *English Premier League* and in Carmichael et al. (2011) as well. Actually, the major proportion of the authors focused their surveys on that competition on the grounds of great

transparency, easy data collection, and belonging to one of the five major European leagues (UEFA, 2020). Fortunately, other economists studied the influence of salaries on team performance for other territories. Both Caruso et al. (2017) and Ferri et al. (2017) examined the case of the Italian *Serie A*. Llorca & Teste (2016) for the French *Ligue 1*, Garcia-del-Barrio & Szymanski (2006) for the Spanish *La Liga*, (Madsen et al., 2018) for the Norwegian and Swedish championships and lastly, Frick (2013) for the German *Bundesliga*. In this latter, a comparative analysis is even made with other European leagues. All of them found out that the relation pay-performance is positively significant.

Summarizing the evidence, it appears that the literature analyzing determinants of football is vast for a decade now. On the other hand, it focuses particularly on "top" five leagues in Europe (England, France, Germany, Italy, and Spain). The investigation intends to extend the documentation to "sub-top" leagues.

### 3 Data

This section is divided into three parts. The first one introduces the first Belgian division and its format. Secondly, the data used in the analysis is presented and discussed, and then the last part shows some descriptive statistics.

#### 3.1 The Jupiler Pro League

The *Jupiler Pro League* is the official name of the first Belgian football division. It was created in 1895, and since 2016, the real denomination is *Jupiler Pro League* or division 1A. It is important to distinguish the JPL from the Pro League since the latter is the department from the Royal Belgian Football Association which manages the professional competitions in Belgium. In other words, the Pro League is a corporation that organizes the structure of the first division (JPL) and the second division (*Proximus League* or Division 1B). Furthermore, they manage the u21 Jupiler Pro League, the national cup from round-of-32 (*Croky Cup*), the Belgian Supercup (battle between the winner of the championship and the winner of the national cup) and finally, the *FIFA* e-sports competition in Belgium. Lastly, the Pro league represents the club's interests, negotiates TV-rights, grants licenses, and is composed of the twenty-four professional clubs in Belgium (Pro League, 2020).

Sixteen teams compete every season in the *Jupiler Pro League*. The season starts at the end of July and finishes in May. A victory is equivalent to three points, a draw to one point and loss, to zero point. Nonetheless, the format of this competition is not traditional. Actually, since the 2009-2010 season, the championship is separated into two stages: the classical phase and the play-offs.

The first stage looks like a regular competition. Every team meets the other fifteen teams twice, altogether thirty games, and content to gain as many points as possible. After that, the worst squad (the sixteenth) is sent into second division for the following year, and the first six go into PO1. Since this current 2019-2020 season, the system of PO2 has changed: from now on, teams who rank seventh to sixteenth in JPL and first to sixth in *Proximus League* are sent into PO2 (even if the sixteenth from the first division is relegated and first from second division is promoted). Afterward, these sixteen teams are split into four groups of four teams. They meet twice. Then, every winner of each group competes in semi-finals (two times again) to access the final. The champion of that final disputes one last game against the team who finished fourth in PO1 to obtain a pass for the second qualifying round in the *Europa League*. Regarding PO1, as previously stated, teams that end the regular season from first to sixth are classified. Every formation faces every opponent twice for a total of ten games. Essential remark, at the beginning of that new tournament, points from stage one are divided by two. The champion is the team that wins the most points at the end of the term. In the case of draw between first and second, the one with higher point during the regular competition is the winner. The second receives a ticket for the third qualifying round in *Champions League* (Pro League, 2020).

The best performing clubs of the regular season compete in play-offs 1, and over the last five years, Belgium had four various champions, whereby PO1 catches most of the attention although PO2 also give a ticket for European qualification. Therefore, it seems reasonable to think that each squad will aim to perform well during the regular season with a view to qualifying for PO1 and being crown best team in Belgium. Nevertheless, this is not the case. Actually, the great majority have lower objectives, and their priorities are to stay in the first division since it means bigger financial payouts and possibly a place in play-offs 1. Only five teams out of the sixteen aspire to be champion every single competition. This small group is the so-called "G5". The G5 is the five teams that have obtained the best sportive results over the last five seasons. It is composed of Anderlecht, Club Brugge, Genk, Gent, and Standard de Liège (Deloitte, 2019). Teams outside of this group belong to the "K11" (Deloitte, 2019).

### 3.2 Data collection and description

The data set contains information about clubs that participated in the *Jupiler Pro League*, from the 2009-2010 season until the 2018-2018 season. This time period has not been taken randomly. Indeed, 2009-2010 marks the implementation of the play-offs system in Belgium. It would have been more complicated to measure team performance in two separate formats. Consequently, this data includes twenty-three

different teams<sup>5</sup> that took part in the competition at least once over the course of ten seasons, for a total of 160 observations. The database is an unbalanced panel<sup>6</sup> due to the system of relegation and promotion. On top of that, out of the database, we will retrieve an extra sample, which is the normal sample, but without the G5. There are 110 observations. In addition, we will create a data set exclusively made up of teams who struggled in play-offs 1 (since their introduction to 2018-2019 season), to analyze this mini-championship. It composed of 60 teams.

It is convenient to add two additional remarks. On the one hand, only eight squads have never experienced the *Proximus League* in the period studied; those are Anderlecht, Club Brugge, Genk, Gent, Standard de Liège (G5), Kortrijk, Lokeren, and Zulte Waregem. On the other hand, during the first season studied, Mouscron went bankrupt, and therefore, formations played twenty-eight games instead of thirty.

At the time of building the database, it was not easy to find data needed for this study because, generally, football details are not freely available. Moreover, some clubs or companies are not always eager to share their own information due to confidential clauses, particularly about wages. Furthermore, comparing to "top-five" European leagues (England, France, Germany, Italy, and Spain), there is a lack of transparency about gates revenue, TV-revenue, and commercial revenue in Belgium. Thankfully, some data were available on websites, periodic, and corporate. Table 1 provides an overview of the variables used and their origins.

Table 1: Variables used in the model and sources

Variable	Definition and source
<b>league_pos</b>	League position. Calculated as below. Retrieved from <i>TransferMarkt</i> . $league\_pos_{i,s}^a = -\log\left(\frac{rank_{i,s}}{23^b - rank_{i,s}}\right) \quad (1)$
<b>pnt_pct</b>	Point percentage. Computed as below. Retrieved from <i>TransferMarkt</i> . $pnt\_pct_{i,s} = \frac{points_{i,s}}{points_s} \quad (2)$

<sup>5</sup>The name of every team plus those who participated in the play-offs 1 are displayed in appendix 1 and 2.

<sup>6</sup>Cases where the individuals, here teams, are not observed over the entire sample period (Baltagi, 2001).

Code	Definition and source
<b>rel_wages</b>	Relative wages. Computed as below. Retrieved from <i>National Bank of Belgium</i> . $rel\_wage_{i,s} = \frac{payroll_{i,s}}{average\_payroll_s} \quad (3)$
<b>rel_rev</b>	Relative revenue. Computed as below. Retrieved from the <i>UEFA</i> and the periodic <i>Le Soir</i> . $rel\_rev_{i,s} = \frac{revenue_{i,s}}{average\_revenue_s} \quad (4)$
<b>net_trans</b>	Net transfers which represents the difference between transfers paid and transfers received. Retrieved from <i>TransferMarkt</i> .
<b>av_age</b>	Average age of the team. Retrieved from <i>TransferMarkt</i> .
<b>foreigners_prop</b>	Foreigners proportion. Computation as below. Retrieved from <i>TransferMarkt</i> . $foreigners\_prop_{i,s} = \frac{foreigners\_players_{i,s}}{roster_{i,s}} \quad (5)$
<b>cup_games</b>	Number of games played in the national cup. Retrieved from <i>TransferMarkt</i> .
<b>euro_games</b>	Number of games played in CL or EL. Retrieved from <i>TransferMarkt</i> .
<b>roster</b>	Total number of players in the squad. Retrieved from <i>TransferMarkt</i> .
<b>filling_rate</b>	Filling rate. Computation as below. Retrieved from <i>European Football Statistics</i> . $filling\_rate_{i,s} = \frac{average\_attendance_{i,s}}{stadium\_capacity_{i,s}} \quad (6)$

<sup>a</sup> i indicates the ith team and s indicates sth season

<sup>b</sup> total number of teams in the sample (adjusted for other samples)

The huge majority of the data come from *Transfertmarkt.com* and *EuropeanFootballStatistics.com*. Both websites are well-known and often used by diverse authors as Frick (2013), Kuper and Szymanski (2018), Madsen et al., (2018), and the company Deloitte (2019). This is because it offers plenty of relevant data

on players, teams, clubs, and competitions in soccer. Thus, the two pages are trustworthy.

League position and point percentage are proxy variables employed to measure seasonal success, and hence, they can be deemed as our dependent variables. A big part of the empirical reviews uses either league position (Hall et al., 2002; Franck & Nüesch, 2011) or point percentage (Carmichael et al., 2011; Yamamura, 2015) or both (Caruso et al., 2017) when trying to gauge performance. Those variables have the advantage to avoid the possibility of randomness (Szymanski & Smith 1997) and take into consideration the gap between top, middle, and low teams. The higher they are, the better the team performance is.

This first independent variable, relative wages, is appropriate to the extent that this variable has always been subject to research (Szymanski & Smith, 1997; Simmons & Forrest, 2002; Frick, 2013; Llorca & Teste, 2016). Concerning details about team payroll, every they are derived from *The National Bank of Belgium* (2020), where annual accounts are published on a yearly basis. Although figures include wages for all the staff, footballers' salaries represent the vast majority and can be considered as relevant.

Then, this thesis incorporates five other regressors: European<sup>7</sup> and national cup games as suggested by Picazo-Tadeo & González-Gómez (2010), the average age, by Bucciol et al. (2014), number of players in the squad, by (Caruso et al., 2017), and ultimately, foreigner number in the team (Bachan et al., 2014).

Next, with the purpose of having as much control as possible, three more explanatory variables are added: filling rate, net transfers, and relative total revenue. Total revenue is the sum of four incomes from the ith club: TV and gates revenue, commercial incomes, and European rewards. The first comes from the Belgian periodic *Le Soir* (2020) and the remainder of a combination from each *UEFA* annual reports and payments (UEFA, 2009-2019).

Lastly, wages, total revenue, and transfers are measured in millions of euros. They were converted from nominal to real figures (relative to 2013) by using the consumer price index provided by the *Belgium Institute of Statistics* (Stabel, 2020). Furthermore, information about 2009, 2010, and 2013 was not displayed. To handle this issue, we took the inflation rate from *StatBel* and deflated numbers with 2014 as the reference year.

---

<sup>7</sup>European matches also contain qualifying matches.

### 3.3 Descriptive statistics

#### 3.3.1 Summary statistics

Table 2 illustrates the summary statistics for the variables<sup>8</sup>. Each team is treated as an individual over the period. We can notice that the number of observations is 159 and not 160. This is due to the fact that in season 2009-2010, the club Mouscron went bankrupt and offered no data.

The summary statistics table gives an interesting preliminary analysis. If we first focus our attention on the internal face of the JPL, it is a rising stars competition and depends mainly on foreign players to stay competitive against national and continental competitions. Indeed, an ordinary team in the sample, over the last ten seasons, is 23.6 years old while foreigner proportion in the squad is on average 60%. Regarding transfers, because the average of net transfers is negative, it proves that Belgian clubs try to buy players and sell them later with a plus-value to make a profit. We can assume that those players come from "low" European leagues or non-European leagues and thus, require lower wages and transfer fees. In addition, on average, attendance is relatively correct where in every game, almost 70% of the stadium is full-filled. This number can even rise to 99% during derbies or crucial games. One last point to emphasize is that some crews are over listed. We learn that sometimes, the roster attains a maximum of 53 players. In other words, more than four starting eleven that can begin a game.

Afterward, when we take a look at the external point of view, Belgian teams do not perform well in European tournaments. Since 2009, the usual path is to play maximum three games. So, in general, clubs do not pass group stages.

#### 3.3.2 Team payroll and team performance

In the scatter plots from figure 1, when we use league position as a dependent variable. We can observe a positive linear relationship between salaries and seasonal success. In spite of few outliers, the results seem in line with the existing literature.

The upward trend when using another dependent variable, i.e., point percentage, tends to be very similar in comparison with figure 1 in appendix 5.

Taking all the incomes jointly, the trend is not as bright as in figure 1. We observe an ambiguous relationship between ranking and total revenue.

---

<sup>8</sup>More descriptive statistics are available in appendix 3, 4.

Table 2: Summary statistics

Variable	N	Mean	Std. Dev.	Min.	Max.
league_pos	159	0.309	0.469	-0.359	1.342
pnt_pct	159	0.452	0.145	0	0.767
rel_wages	159	1.006	0.801	0.051	3.973
rel_rev	159	1.021	0.674	0.404	2.965
net_trans	159	-2.148	6.109	-33.276	12.210
av_age	159	23.624	1.030	21.700	26.800
foreigners_prop	159	0.539	0.118	0.219	0.786
cup_games	159	3.214	1.897	1	8
euro_games	159	2.516	4.359	0	18
roster	159	35.497	5.202	26	53
filling_rate	159	0.673	0.191	0.227	0.996

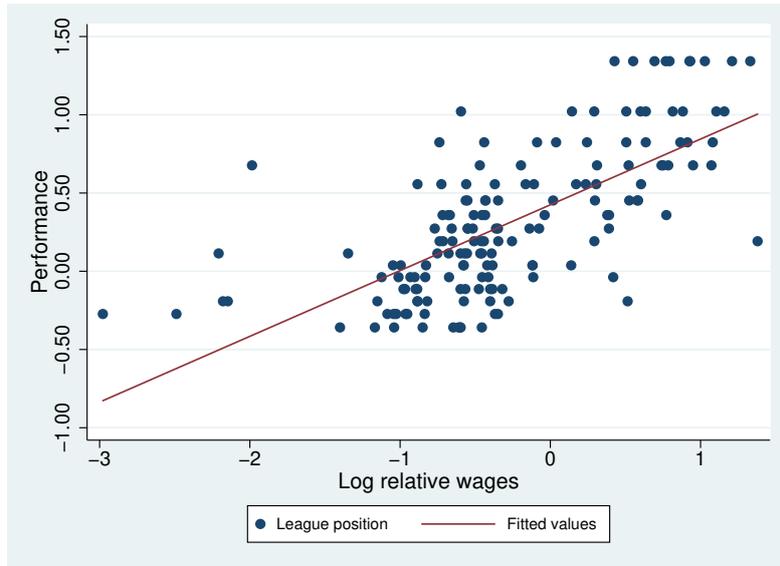


Figure 1: Relation team payroll on team performance (1).

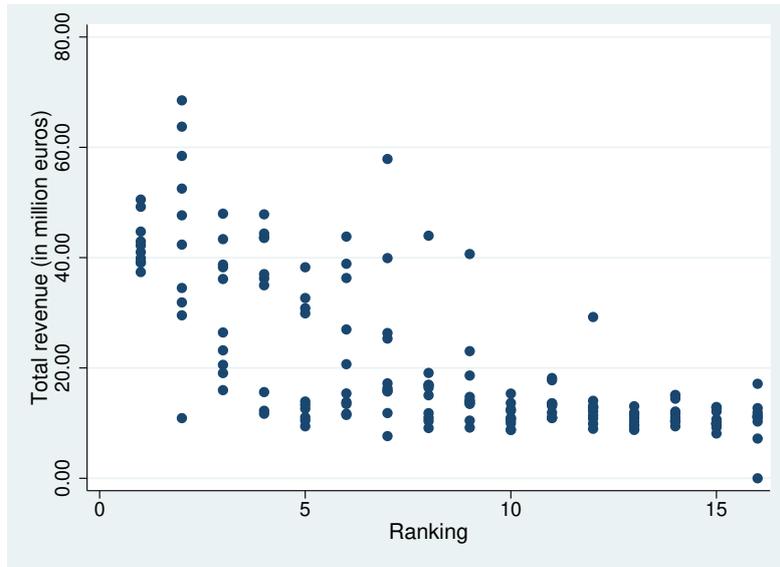


Figure 2: Relation revenue and ranking.

Source: Own data analysis drawn from Stata 14 (both figures).

## 4 Empirical approach

This section aims at presenting the methods applied in order to discover which variables might have an influence on team performance.

### 4.1 Model

As presented in section 3, the data set includes information about twenty-three teams on ten seasons. Hence, we have an unbalanced panel data of 160 observations. This permits to use panel data analyses, which is very convenient since it provides more efficient and consistent estimators than simple cross-sectional data. It also handles the problem of endogeneity (Wooldridge, 2016). Fixed Effects estimator (Within estimator; FE) or Random Effects (RE) estimator are two common methods when estimating panel data. A Fixed effect estimator is very handy because it removes time fixed effects; however, at the cost of decreasing the number degrees of freedom and removing regressors that do not vary over time. Random effect estimator does not apply this mechanism and tends to be more attractive but it is based on two assumptions: the explanatory variables are drawn randomly from a population and uncorrelated with unobserved factors (Baltagi, 2001), which is a very strong assumption.

When analyzing the model, it exhibits features to that of the Fixed-Effects model because we suppose that we are in the presence of individual-specific effects. Indeed, every team has its own individual unobserved characteristics, such as pressure from the club's city if fan interest in football is huge, its own history, its own way of spending money or even its tendency to promote players from their academies, etc. These elements can impact the estimated parameters, and thus, it is crucial to control for them to decrease the omitted bias (adding team dummies). Similarly, time unobserved variables, such as league reforms, will also be under control by adding seasonal dummies variables. In other words, we monitor seasonal changes that can impact final rankings and clubs with financial problems, possibly resulting in them giving up the competition or new better training facilities available for a squad.

Eventually, relative wages and relative revenue are elevated in logarithms forms (not net transfers due to negative observations) and the square of average age and roster are taken. Logarithms forms compared to level forms have several advantages because elasticity interpretation is possible. This is important since variables tend to have a normal distribution and decrease the effect of the outliers. On the other hand, square variables offer the possibility to interpret decreasing or increasing marginal effects (Wooldridge, 2016).

The model to estimate, where index "i" represents the ith team ( $i = 1, \dots, 23$ ) and index "s" represents the sth season ( $s = 1, \dots, 10$ ), is presented as hereby:

$$Y_{i,s} = \beta_0 + \beta_1 \text{lrel\_wages}_{i,s} + \beta_2 \text{lrel\_rev}_{i,s} + \beta_3 \text{av\_age}_{i,s} + \beta_4 \text{net\_trans}_{i,s} + \beta_5 \text{sqav\_age}_{i,s} + \beta_6 \text{foreigners\_prop}_{i,s} + \beta_7 \text{cup\_games}_{i,s} + \beta_8 \text{euro\_games}_{i,s} + \beta_9 \text{roster}_{i,s} + \beta_{10} \text{sqroster}_{i,s} + \beta_{11} \text{filling\_rate}_{i,s} + (\delta_1 \sum_{i=2}^{23} \text{team}_i + \delta_2 \sum_{s=10}^{18} \text{season}_s) + \alpha_i + \gamma_s + \epsilon_{i,s}$$

Where,  $Y_{i,s}$  equals our dependent variables,  $\beta_0$  is the constant, where  $\alpha_i$  and  $\gamma_s$  are respectively teams and seasons invariant terms, and  $\epsilon_{i,s}$  is the idiosyncratic error term.

The first estimator employed to run the model is the conventional Pooled OLS. This estimator is very useful because it multiplies the number of individuals by the time period and thus, increasing the sample size and the statistical power (Wooldridge, 2016).

Next, a Hausman test is performed to establish rather RE or FE should be implemented. This test basically checks whether the unique errors are correlated with the regressors. In this case, the null hypothesis stated that differences in the coefficients is not systematic, is not rejected (p-value = 0.233). So here, RE and RE are close enough that it does not matter which one to choose (Wooldridge, 2016). Since we are in the presence of fixed effects, using a RE estimator will lead to unfair results (Cameron & Trivedi, 2012). So, in this situation, we will perform the Within estimator because it allows for some correlation between constant unobserved factors and regressors. Following that, the Wald test indicates that no seasonal dummies are not needed when estimating by FE (p-value = 0.946).

Lastly, we ran a Fisher test to verify the pertinence of considering all variables together. Its p-value (0.000) strongly rejects the null hypothesis and indicates that our explanatory variables are relevant to explain our model. On top of that, a Brush-Pagan test and a serial correlation test were performed to detect eventual heteroskedasticity (non-constant variance) or autocorrelation issues. If no attention is given to that, results are still unbiased but no more efficient because a variable is correlated to the variance or with past events. In both cases, we reject the assumption on constant variance and no autocorrelation at the 0.05 level. To drawn reliable conclusions and handle the problem, all the standard errors in the regressions are adjusted for clusters in teams. It enables to deal with non-constant variance and serial correlations issues. In addition, in order to produce consistent and unbiased estimators, we need to assume the strict exogeneity assumption, i.e., the explanatory variables are uncorrelated with the idiosyncratic error term ( $\epsilon_{i,s}$ ) in all time periods (Wooldridge, 2016)<sup>9</sup>.

---

<sup>9</sup>Actually, there are three more assumptions in Wooldridge: linearity in parameters, random

## 4.2 Endogeneity check

So far, to obtain unbiased and consistent estimators, we had to make the assumption of strict exogeneity. Nonetheless, this assumption is a strong assumption because it requires that past, present (and future) events must not have any influence on current realization. In the case of our model, this is probably unrealistic. It is reasonable to suspect that past performance probably determines current salaries and revenue. Most of the time, if a team performs nicely during the championship, managers are likely to renegotiate players' contracts for the following year because they played well and avoid a transfer to a competitor. Of course, at the hour of bargaining, players would want at least, more than what they already have. Besides, when a team is competing in the first roles, it attracts a higher hype, and it will qualify for next European tournaments, both involve larger incomes. Therefore, we have variables that suffer unfortunately from endogeneity<sup>10</sup> with past realizations. A solution to control this issue is to find an instrumental variable (if data availability and this IV must be exogenous and correlated with the endogenous variable) and then, to estimate by 2SLS or add a lagged variable. In this situation, the second option is selected due to the lack of finding a good instrumental variable. However, the problem when applying this method is that we pass from a static panel data to a dynamic panel data (which means interaction of past realizations) and the standard estimators like Pooled OLS, RE estimates or FE estimates are no more valid because they lie on the assumption discussed<sup>11</sup> just above (Baltagi, 2001). Therefore, we need other estimation methods. Caruso et al. (2017) and Yamamura (2015) proposed the System-GMM estimator to fix this matter. This estimator is designed so that in the presence of dynamic panel, regressors might be endogenous, period T is smaller than individuals N, presence of fixed effects, and presence of heteroskedasticity or serial correlation (Roodman, 2009). Conditions that we perfectly meet.

In reality, we can find two GMM estimator, the Difference GMM (from Arellano and Bond (1991)) or the System GMM (from Blundell and Bond (1998)). Both correct endogeneity, but they implement two distinct techniques: "first difference" and "orthogonal deviations". Differencing subtracts the variable in the last period observation to the current one, so  $\Delta y = y_1 - y_2$ . The problem with this method is that when we are in the presence of an unbalanced panel, the gap between observations can increase and thus, decrease the statistical power. Plus, the lagged

---

sampling and no perfect collinearity and variability across variables, but all of them are respected in the sample.

<sup>10</sup>According to Cameron & Trivedi (2012), a simple definition of endogeneity is an explanatory variable correlated with an unobserved variable.

<sup>11</sup>Indeed, they assume strict exogeneity. If a lagged variable is included, it implies a correlation with past values.

variable can still be correlated with the error term because with this method, we always need previous observations of it. Orthogonality, however, subtracts current variable observation by the average of future observations (possible until n-1), so  $\Delta y = y_1 - \bar{y}_{future}$ . In addition, to removing constant fixed effects, the data loss is minimized, and, in the computation, no lagged variable is necessary and constant term can be kept. Hence, the system GMM is more desirable.

The new model (with a lagged dependent variable) to estimate, where again index "i" represents the ith team ( $i = 1, \dots, 23$ ) and index "s" represents the sth season ( $s = 1, \dots, 10$ ), is:

$$\Delta Y_{i,s} = \beta_0 + \beta_1 \Delta lrel\_wages_{i,s} + \beta_2 \Delta lrel\_rev_{i,s} + \beta_3 \Delta net\_trans_{i,s} + \beta_4 \Delta av\_age_{i,s} + \beta_5 \Delta sqav\_age_{i,s} + \beta_6 \Delta foreigners\_prop_{i,s} + \beta_7 \Delta cup\_games_{i,s} + \beta_8 \Delta euro\_games_{i,s} + \beta_9 \Delta roster_{i,s} + \beta_{10} \Delta sqroster_{i,s} + \beta_{10} \Delta filling\_rate_{i,s} + \beta_{11} \Delta lag Y_{i,s} + \Delta \epsilon_{i,s}$$

Where again,  $Y_{i,s}$  equals our two dependent variables,  $\beta_0$  is the constant, and  $\epsilon_{i,s}$  is the idiosyncratic error term ( $\alpha_i$  and  $\gamma_i$  have disappeared thanks to the orthogonality process).

To estimate this equation, we will perform a Two-System GMM, which is more efficient than the One-System GMM (Roodman, 2009). Dependent variables are instrumented with one period lag.

## 5 Results

This section reveals the results of all the regressions. We are going to try to capture the determinants of triumph in Belgium with the normal sample, which corresponds to the whole data set for regular seasons. Then, we look at the robustness of these main results by using another dependent variable and by making some sample adjustments to retrieve two new ones; the first is the usual sample without the G5, and the second equals the competition of play-offs 1 only. All three samples are from the championship 2009-2010 until 2018-2019. The main purpose is to detect if there is any influence from the G5 on K11's success and if winning the championship during the second stage of the competition is due to randomness or no.

## 5.1 Results for the total sample

### 5.1.1 League position as the dependent variable

Table 3 displays the results<sup>12</sup> of the different estimators for the complete sample during the regular season.

The first columns correspond to Pooled OLS with time dummies and the second to Pooled OLS as well, but this time with time and team dummies. At first sight, the model seems to explain the dependent variable nicely because both adjusted  $R^2$  are quite high<sup>13</sup>. Indeed, the regressors explain about 64% and 67% of the variation in league position.

When taking the estimates from (1), the variable relative wages is significant at the 0.05 level. Holding other variables constant, a one percent increase in team payroll increases by 0.231% seasonal performance. Similarly, relative revenue is also positively correlated with ranking and statistically significant when estimated by Pooled OLS. Then, the coefficient of average age is statistically significant at the 0.05 level; meaning that when the average age of a squad increases by one, league position rises by 184.2%. Nevertheless, because of the square variable is also significant with a negative sign, the positive effect has a marginal effect. This means that the return of average age is positive until 24 years old<sup>14</sup>. After that, it becomes negative. With reference to extra European games, surprisingly, the participation in the *Champions League* or the *Europa League* has counter-intuitive sign. We could expect that the more games ones play, the least energy they will have remaining throughout the season is, but in our case, this is the contrary. In other words, one more party in CL or EL boosts team performance by 1.7%, *ceteris paribus*. Lastly, Pooled OLS reveals the filling rate during the season is significant at the 0.10 level. As a result, all things being equal, an additional percentage point in the stadium-filling improves team success by 29%.

Next, when we add team dummies in regression (2), the Pooled OLS tends to be less attractive in view of the fact that only the filling rate variable remains significant with the same sign as in the previous regression. One possible explanation for this incident is probably due to the fact that coefficients in (1) absorbed unseen team characteristics and were possibly biased, however, now that we explicitly control them, this absorption has disappeared. Team dummies capture an outcome that we do not have control in our model and prove that the basic estimator for panel data is a good indicator but can involves misleading results. Hence, deeper examinations are required. Only filling rate variable remains significant and the variation has

---

<sup>12</sup>Additional regressions are available in appendix 6.

<sup>13</sup>The adjusted  $R^2$  is a better measure than the  $R^2$  because the  $R^2$  keeps increasing when adding a new independent variable. The adjusted  $R^2$ , however, gives a penalty for adding new independent variables (Wooldridge, 2016).

<sup>14</sup>The computation of the turning point is  $|\hat{\beta}_1 / 2\hat{\beta}_2| = 23.61$ .

Table 3: Regression table for regular competition (league position as the DV)

	(1)	(2)	(3)	(4)
	Pooled OLS	Pooled OLS	Fixed Effets	Two-System GMM
lrel_wages	<b>0.131**</b> (0.047)	0.054 (0.038)	<b>0.044*</b> (0.026)	<b>0.245**</b> (0.111)
lrel_rev	<b>0.355***</b> (0.111)	0.072 (0.265)	0.098 (0.197)	-0.889 (0.652)
net_trans	-0.000 (0.003)	0.000 (0.003)	0.001 (0.003)	0.007 (0.009)
av_age	<b>1.842**</b> (0.849)	1.544 (1.115)	1.458 (0.877)	-1.209 (7.823)
sqav_age	<b>-0.039**</b> (0.018)	-0.033 (0.024)	-0.031 (0.019)	0.023 (0.165)
foreigners_prop	0.279 (0.256)	0.234 (0.324)	0.248 (0.249)	0.935 (0.667)
cup_games	0.020 (0.012)	0.027 (0.016)	<b>0.026**</b> (0.013)	0.034 (0.034)
euro_games	<b>0.017*</b> (0.010)	0.013 (0.010)	0.014 (0.008)	<b>0.114***</b> (0.038)
roster	-0.028 (0.037)	-0.032 (0.047)	-0.021 (0.034)	0.118 (0.076)
sqroster	0.000 (0.000)	0.000 (0.001)	0.000 (0.000)	<b>-0.002*</b> (0.001)
filling_rate	<b>0.290*</b> (0.144)	<b>0.904*</b> (0.444)	<b>0.783*</b> (0.400)	<b>1.915***</b> (0.585)
L.league_pos	.	.	.	<b>-0.423*</b> (0.222)
_cons	<b>-21.235**</b> (9.898)	-17.401 (12.988)	-16.998 (10.189)	11.713 (92.523)
N	159	159	159	130
adj. R <sup>2</sup> (overall)	0.643	0.665	0.602	.
Season FE	Yes	Yes	No	No
Team FE	No	Yes	Yes	Yes
Ar1 (p-val)	.	.	.	0.015
Ar2 (p-val)	.	.	.	0.113
Hansen (p-val)	.	.	.	0.231

Clustered robust standard errors in parentheses

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

Source: Own data analysis drawn from Stata 14

enlarged.

Afterward, in estimation (3), the adjusted  $R^2$  has reduced to 60%. However, this is not a big issue given that the within estimation deletes constant unobserved terms (as well as observed terms) and hence, is more reliable. The first striking example is with the variable relative wages, which gets back significant at the 0.10 level. The Fixed effects estimate also shows that variables, such as relative revenue and age with its square, are no more significant comparing to pooled OLS without control for teams and confirms estimates in (2). In contrast, significant relation exists now between sports performance and the number of games played in the domestic cup. This differs from columns (1) and (2). Despite of this new significant variable, the analysis goes in the same way that the one for additional European parties<sup>15</sup>. All else being equal, one more game in the national Belgian cup improves team performance by 2.6%. Again, the variable filling rate remains statistically significant. When keeping other factors constant, a one percentage point increase in filling rate enhances success by 78.3%.

Finally, in the column 4, using Two-System GMM, we are able to control for potential dynamic settings from past realizations as described in section 4. As table 3 indicates, the addition of the lagged dependent variable is meaningful and strengthen its validity. The auto correlation's p-value of order one is small enough to reject the null hypothesis of the absence of auto correlation for order one, but the null hypothesis for order two cannot be rejected. Therefore, the estimator is consistent when no second-order serial correlation is present. More so, the Hansen test does not reject the null hypothesis stating overall validity of the instrumental variables used. Now, coefficients have changed in comparison with Fixed effect. The variable relative wages remains significant, but now at 0.05 level, and its magnitude has increased. Here, one percent growth in relative wages boosts league position by 0.245%, *ceteris paribus*. Once more, the impact of additional matches during the soccer year championship is statistically significant and rise ranking position by 11.4% when equal terms. Finally, as expected, filling rate remains strongly positively correlated to our dependent variable.

### 5.1.2 Point percentage as the dependent variable

The second explained variable, elevated in logarithm form, aims at checking the veracity of the findings in 5.1.1. Table 4 reports the results.

Pooled OLS coefficients without and with team dummies are detailed in column (1) and (2). The  $R^2$  are respectively a 5% and 7% lower than those with league position as a dependent variable, so the loss is very low. Now, the information

---

<sup>15</sup>In regression (1), supplementary European games were positively related to the dependent variable.

contained in point percentage is explained at 59% by the regressors in the pooled OLS benchmarks.

In regression (1), all variables, except extra European games, are statistically significant as in table 3; the filling rate at the 0.01 level, relative wages, and relative revenue at the 0.05 level and average at the 0.10 level. Although all have a positive effect on point percentage, their magnitude has now declined (all other things being equal). The turning point for average age equals now 23 years old, which is one year younger. The filling rate variable is the only exception where the influence has improved by 0.143. A one more percentage point in filling rate augments team performance by 43.3%, *ceteris paribus*.

Then, when taking Pooled OLS in (2), it yields exactly the same conclusion as regression (2) previously displayed. The variable observing attendance during the season is meaningful, but this time, the coefficient becomes smaller.

The third regression estimates the equation by Fixed Effects. The adjusted  $R^2$  decreases again by 1% to equal 59% here. Regarding the betas estimated, although the Within estimation offers the same outcomes as column (3) in table 3 with a lower extent, there are two notable exceptions to highlight. Indeed, relative wages has no more substantial evidence, and foreigners proportion is unexpectedly significant at the 0.10 level when constant error term is deleted, and this counter our previous conclusions. Everything else being equal, an enlargement of one percentage point for non-native players in a squad increases seasonal achievement by 38.1%.

Ultimately, a Two-System GMM is performed in the last column to add control for possible past influence. In this case, the lagged dependent variable is no longer significant but still allows for past linkage. Once more, autocorrelation's p-values reject the null hypothesis of no auto correlation of order one, meanwhile autocorrelation of order two cannot be rejected. Once more, the Hansen test does not reject the null hypothesis stating that the overall validity of the instruments. In this regression, the betas show that relative wages regains significance at the 0.05 level in comparison with (2) and (3), and are in agreement with the first GMM estimator implemented in this thesis. One again, its effect is smaller and implies an increase in seasonal success by 0.169%, with all other factors being equal. Contrary to table 3, going through European competition and fill-in stadiums are no more relevant.

### 5.1.3 Comparison

Regressions, using two different proxies variables for measuring squad performance through the regular competition, enhanced three main independent variables: relative wages, additional games, and the filling rate.

Firstly, growth in team payroll has a significant impact on positive seasonal

Table 4: Regression table for regular competition (point percentage as DV)

	(1)	(2)	(3)	(4)
	Pooled OLS	Pooled OLS	Fixed Effets	Two-System GMM
lrel_wages	<b>0.096**</b> (0.041)	0.061 (0.045)	0.044 (0.030)	<b>0.169**</b> (0.069)
lrel_rev	<b>0.185**</b> (0.076)	0.084 (0.187)	0.127 (0.157)	0.221 (0.528)
trans_net	-0.001 (0.002)	-0.001 (0.002)	-0.001 (0.002)	0.003 (0.003)
av_age	<b>1.092*</b> (0.592)	1.283 (0.840)	1.255 (0.733)	-0.070 (1.076)
sqav_age	<b>-0.024*</b> (0.013)	-0.028 (0.018)	<b>-0.027*</b> (0.016)	0.001 (0.023)
foreigners_prop	0.300 (0.209)	0.327 (0.270)	<b>0.381*</b> (0.196)	0.399 (0.507)
cup_games	0.012 (0.009)	0.019 (0.011)	<b>0.018*</b> (0.009)	-0.002 (0.016)
euro_games	0.006 (0.005)	0.004 (0.005)	0.004 (0.004)	0.001 (0.034)
roster	-0.037 (0.026)	-0.034 (0.037)	-0.020 (0.027)	0.018 (0.046)
sqroster	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	-0.000 (0.001)
filling_rate	<b>0.433***</b> (0.072)	<b>0.801**</b> (0.325)	<b>0.669**</b> (0.304)	0.579 (0.705)
L.lpnt_pct	.	.	.	-0.027 (0.347)
_cons	<b>-13.066*</b> (6.733)	-15.352 (9.694)	<b>-15.341*</b> (8.389)	-0.389 (12.376)
<i>N</i>	159	159	159	130
adj. <i>R</i> <sup>2</sup> (overall)	0.592	0.592	0.586	.
Season FE	Yes	Yes	No	No
Team FE	No	Yes	Yes	Yes
Ar1 (p-val)	.	.	.	0.065
Ar2 (p-val)	.	.	.	0.682
Hansen (p-val)	.	.	.	0.331

Clustered robust standard errors in parentheses

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

Source: Own data analysis drawn from Stata 14

outcome in the *Jupiler Pro League* and confirms section 3.3.2 and the empirical reviews regarding this topic (Simmons & Forrest, 2002; Carmichael et al., 2011; Llorca & Teste, 2016; Ferri et al., 2017; Kuper & Szymanski 2018; Madsen et al., 2018). It is even more convincing when the dynamic setting is implemented, and team fixed effect is erased. In other words, larger team payroll increases success. Plausible justifications for the importance of salary is primarily due to the format of the competition. The championship in Belgium is an open league and a free football market, unlike in the United States where the members of the league never change, and salary cap exists, for instance. Consequently, the system of relegation and promotion encourages teams to invest more in remuneration in order to attract better players and to remain in the first division.

Secondly, as this was unexpected, crews who play extra games during the season (in European or Belgian cups), will achieve better ranking at the end of the regular season. This finding matches with Picazo-Tadeo & González-Gómez<sup>16</sup> (2010). A reason for this might be that when players keep qualifying for next round in European tournaments, they are more motivated to win these special trophies and locate necessary resources to stay competitive in participating competitions. On the contrary, it is also essential noting that, half of clubs qualifying for European participation, or clubs that reached the final of the Belgium cup, are clubs from the G5 and thus, can sometimes mislead the results because they compete every year for the title.

Thirdly, as anticipated, the more people coming to the stadium, the better team's performance is. Sometimes, the crowd is so powerful that it allows players to go the extra mile and win games. Thus, the role of the "twelfth man" is very important across seasons. It motivates players, while also meant as a form of pressure on them to stay at the top.

On the contrary, neither relative revenue nor net transfers have substantial effects on our dependent variable, except with Pooled OLS in (1) for relative revenue, but as explained previously in section, this has drawbacks related to controls of unobserved effects. As a result, teams do not have to generate tons of earnings or waste money regarding players' moves. This is puzzling since clubs are paying more and more for transfers fees every summer. However, Kuper & Szymanski (2018) suppose that incomes might have indirect relationship with seasonal achievement. Certainly, clubs with higher incomes do always win leagues, but do have greater influence to spend more on team payroll and thereby, increase their probability to gain for example. In all cases, there is no direct relationship.

In terms of average age and roster, they do not play any role in the way of

---

<sup>16</sup>In their paper, they emphasized the necessity of integrating additional games when estimating and that playing more parties throughout the championship is mandatorily detrimental.

winning the Belgian league, contradicting Tillou and Liarte (2012)<sup>17</sup> or Caruso et al. (2017)<sup>18</sup> who demonstrated significant outcomes, especially in the System GMM framework. This means that having more players in the squad is not necessarily an evidence of more talent availability. Nevertheless, it is worth mentioning that the first ones based their outcomes on the National team of Brazil and not a league, and the second ones worked on a different league (*Serie A* in Italy) for a different time period studied (from 2001-2002 to 2014-2014).

Lastly, in spite of, on average, half of the squad coming from abroad, having more foreigner players is no a warranty of better team performance and confirms what Royuela & Gásquez (2019) and Bachan et al. (2014)<sup>19</sup> analysis.

## 5.2 Results for K11 sample

This adjusted sample is devoted to investigate on whatsoever influence from the G5 on teams considered as "smaller". The data set is now an unbalanced panel of 110 observations with eighteen different clubs<sup>20</sup>. All the tests effectuated for the total sample are still valid, whereas variables are jointly significant and all the standard errors in the regressions are one more time, adjusted for clusters in teams.

### 5.2.1 League position as the dependent variable

Table 5 exposes the results for the regular season, but without the G5. The same estimators as earlier (Pooled OLS with season dummies, Pooled OLS with season and team dummies and Fixed effects estimator) are executed except the System-GMM estimator, sicne to be implemented, it requires the number of instruments to be lower or equal to the number of groups. Otherwise, the model suffers from overidentification (Roodman, 2009). In current case, the number of instruments is twenty-two (as with the complete sample) but the number of groups is seventeen<sup>21</sup>. Hence, we will work only in a static setting.

Primarily, when we take a look at the adjusted  $R^2$ , the independent variables account for 33% of the variation in league position in column (1) and slightly improves until 37% in column (2) and 38% in column (3). The power of the

---

<sup>17</sup>They showed that the best mix for a team is to have as much young players as old players

<sup>18</sup>In their paper, they found out significant relationship between age and roster with sport performance.

<sup>19</sup>They found no evidence of better results within league when teams own large foreigners players.

<sup>20</sup>In reality, the number of observation is 109 and number of different clubs is seventeen because Mouscron went into bankruptcy during the season 2009-2010.

<sup>21</sup>Actually, the number of instruments is the instrumented variables and the groups are the number of clubs in the sample.

Table 5: Regression table for K11 sample

	(1)	(2)	(3)
	Pooled OLS	Pooled OLS	Fixed Effets
lrel_wages	<b>0.130**</b> (0.050)	0.053 (0.056)	0.033 (0.053)
lrel_rev	<b>0.839***</b> (0.222)	<b>1.713*</b> (0.813)	<b>1.228**</b> (0.517)
trans_net	0.009 (0.009)	0.009 (0.013)	0.008 (0.010)
av_age	1.282 (1.474)	1.685 (1.484)	1.310 (1.266)
sqav_age	-0.028 (0.031)	-0.036 (0.032)	-0.028 (0.027)
foreigners_prop	<b>0.785*</b> (0.384)	0.691 (0.632)	0.432 (0.449)
cup_games	0.007 (0.028)	-0.012 (0.050)	-0.001 (0.033)
euro_games	0.026 (0.016)	-0.024 (0.020)	-0.006 (0.018)
roster	-0.098 (0.069)	-0.056 (0.095)	-0.066 (0.056)
sqroster	0.001 (0.001)	0.000 (0.001)	0.000 (0.001)
filling_rate	<b>0.299*</b> (0.169)	-0.389 (1.042)	0.012 (0.749)
_cons	-12.820 (16.728)	-17.622 (16.661)	-13.298 (14.605)
<i>N</i>	109	109	109
adj. <i>R</i> <sup>2</sup> (overall)	0.332	0.367	0.384
Season FE	Yes	Yes	No
Team FE	No	Yes	Yes

Clustered robust standard errors in parentheses

\*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$

Source: Own data analysis drawn from Stata 14

regressors have almost declined by half. Therefore, the model explains much less the actual sample than the full sample.

The first column enables an insight of the model estimated by Pooled OLS with seasonal dummies. The three variables; relative wages, relative revenue, and filling rate are statistically significant as they were in the total sample. Holding all factors constant, one percent increase in relative salaries and relative incomes raises team success by respectively 0.13% and 0.839%, meanwhile one additional point percentage in filling rate boosts the dependent variable by 29.9%. Conversely, a new variable, the amount of foreigners in the squad, is significant at the 0.10 level and positively associated to ranking.

The second regression replicates the same analysis as in (1) but with more control. Now, no variables have an effect on seasonal achievement. Relative revenue remains meaningful (lower level) at the 0.10 level.

The last regression uses Within estimate, to which more attention should be given due to its properties of dealing with unobserved characteristics. It confirms the conclusions provided in Pooled OLS (2). However, the magnitude of the coefficient is lower now even if the significance level and has passed from 0.10 to 0.05. Everything else equals, one percent growth in relative revenue improves league standing by 1.228%.

### 5.2.2 Comparison with total sample

The sub-sample reveals new conclusions in comparison with that of all squads. It is discovered that paying higher wages is no longer indispensable, but total revenue plays a major role in the way of achieving good season. A possible reason for this is due to the majority of team objectives being not necessarily to win the competition, but to remain in the division. Many teams are struggling every year to achieve this goal with less resources to hire players considered as more "skillful" because of salaries. Thus, in the long run, generating larger incomes helps clubs to become more competitive because they are able to establish better training facilities to practice in better conditions, and in some cases, develop their youth academy to form better players or improve their coaches' formations.

We conclude that in order to leave the bottom of the ranking, higher recipes are needed to develop the club and increase .

## 5.3 Results for play-offs 1 sample

This new transformation sample is exclusively composed by twelve different teams who were able to qualify for play-offs 1 since season 2009-2010. The data set includes 60 observations. We remind that only the six first teams at the end of the

regular season obtain their ticket to access this mini-championship and the points gained are divided by two before starting.

### 5.3.1 League position as the dependent variable

To estimate this new sample, we will add one more regressors: initial league position which corresponds to the rank at the end of the regular season. We perform Pooled OLS and Fixed effects estimators. However, the System-GMM approach is not executed for the same reasons exposed as in section 5.2. Table 7 shows the estimated coefficients with adjusted clusters in teams.

The first regression from column (1) reports an adjusted  $R^2$  of 53%. Two variables are significant: the first one, at the 0.01 level, is not surprisingly the final position at the end of the regular competition. Holding others factors constant, higher initial league position when starting this second stage of the competition is positively correlated to success in play-offs 1. The second variable, at the 0.10 level, is extra continental games. Here, one more party in CL or EL penalizes sport performance by 2.6%, all else being equal.

Pooled OLS with both team and seasonal dummies offers on the one hand the same results as in (1), with stronger coefficients but with a lower  $R^2$  (45%). On the other other, the regression proposes two additional significant variables at the 0.10 level. These two are relative wages and foreigners proportion. Larger team payroll has a counter intuitive sign.

Indeed here, it is expected to decrease success during plays-offs 1 under equal conditions.

Lastly, the Fixed effects estimator, which is the most efficient, yields similar outcomes as in (1) and erases the effect of relative wages and foreigners proportion from (2). The  $R^2$  proposed re increases by 2%. One percent growth in initial league position before starting the play-offs 1 improves team performance in this latter by 0.913%. Similarly, one more game in European cup is expected to decrease final ranking during play-offs 1 to by 3.1%, *ceteris paribus*.

### 5.3.2 Comparison with the total sample

Primarily, we found out that finishing the classical stage of the competition with a higher ranking is decisive to improve triumph in the play-offs 1 contest. This means that performing well during the regular season provides a substantial advantage when competing during this mini-championship to be crown champion. Team salaries is no more significant but it is evident that it has an indirect effect because it permits to increase league position before the regular season.

Regarding the effect of extra games (particularly European games) now, it is deleterious when participating in PO1. This means that when the question of

Table 6: Regression table for play-offs 1 sample

	(1)	(2)	(3)
	Pooled OLS	Pooled OLS	Fixed Effetes
ini_league_pos	<b>0.825***</b> (0.157)	<b>0.947***</b> (0.202)	<b>0.913***</b> (0.161)
lrel_wages	0.063 (0.110)	<b>-0.235*</b> (0.107)	-0.189 (0.108)
lrel_rev	0.241 (0.156)	0.286 (0.313)	0.404 (0.237)
trans_net	-0.002 (0.004)	-0.002 (0.004)	-0.001 (0.003)
av_age	-0.630 (1.061)	-0.168 (2.474)	0.161 (1.675)
sqav_age	0.012 (0.023)	0.002 (0.053)	-0.004 (0.036)
foreigners_prop	0.158 (0.446)	<b>-0.776*</b> (0.387)	-0.620 (0.478)
cup_games	0.005 (0.019)	0.003 (0.028)	0.003 (0.021)
euro_games	<b>-0.026*</b> (0.012)	<b>-0.031*</b> (0.015)	<b>-0.031**</b> (0.012)
roster	-0.007 (0.061)	-0.092 (0.125)	-0.077 (0.113)
sqroster	-0.000 (0.001)	0.001 (0.002)	0.001 (0.001)
filling_rate	0.114 (0.325)	-0.133 (0.993)	-0.216 (0.753)
_cons	8.218 (12.399)	5.418 (29.155)	0.681 (19.486)
<i>N</i>	60	60	60
adj. $R^2$	0.525	0.451	0.469
Season FE	Yes	Yes	No
Team FE	No	Yes	Yes

Clustered robust standard errors in parentheses

\*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$

Source: Own data analysis drawn from Stata 14

"which competition favour" arises only if teams access the final-six.

More surprisingly, the effect of more attendance has disappeared. A plausible explanation is that these high-stakes games attract a lot of people and thus, almost all stadiums are full in order to support their favorite clubs.

## 6 Policy recommendations

Based on the investigations made through this thesis, several policy recommendations can be mentioned.

First and foremost, relative wages appear to be a very good incentive to promote victory across seasons and improve final ranking before starting the playoffs one contest. Football directors should focus their attention on current team payroll instead of buying players for exorbitant prices, for instance. Obviously, nowadays, release clauses exist, and consequently, it is harder to seduce new players with attractive salaries since most clubs will typically try to do the same. Nevertheless, there are ways around this. The first one is waiting until the footballer becomes a free agent, as exposed in the introduction. The second one is to negotiate with them when it remains his one-year contract with his current club because the player knows that if his current club does not accept his conditions, he will be able to move next year for free. Thus, managers should not hesitate to spend money on wages but bargain as much as possible for players' prices because the former helps to succeed in the long term. Clubs that are least flexible for high wages, it is preferable to generate financial incomes. The new TV-deal, if split equally among teams, as in England for example, might enlarge competition by helping financially teams considered as "small".

Afterward, additional games outside of the championship match must not be neglected. During the regular season, because it is not detrimental at all. The last ten seasons, few Belgian squads have reached at least the quarter-finals of a tournament organized by the UEFA since they appear to concentrate more on domestic games. This, however, is not the best behavior to adopt if they want to broaden their possibilities to win trophies. We have seen that the more crews were advancing into the knock-out stages, the more they get the motivation to realize history because this phenomenon probably happens once or twice in a career. Moreover, we also discovered statistical evidence that teams who went the extra mile in Belgium cup performed better the same season throughout the regular competition. Although supplementary games do not spoilsport performance during the regular season, when players succeed in obtaining their qualifications for the play-offs 1, it becomes detrimental. The question of which competition we should focus on arises at this moment.

Next, the presence of supporters across the season is vital in improving team

achievement. Despite that this is not something new and average attendance is relatively correct (67.3%), Belgian clubs are still far from their German and Dutch neighborhoods - *Bundesliga* and the Eredivisie reach respectively 91% and 88% (European Leagues, 2018). It highlights the crucial condition of keeping constant supporters coming to the stadium. Possible solutions to improve fan base is either include considering better costumer-relationship and customer experience schemes with the help of marketing and social media, or building new arenas, as Kuper Szymanski (2018) proposed.

As another relevant point, we concluded that outlander players do not contribute particularly to bringing triumph. Rather than wasting money only to recruit footballers coming outside of the Belgian's borders, a wise decision would be to split those investments into recruitment and academy development.

Finally, this thesis recommends much more data availability. The *Jupiler Pro League* should have accessible data accessible regarding the financial and nonfinancial activities of its professional affiliates, as it is done in Italy, Germany, Spain, the UK, and in other sports such as in NBA or MLB. Few investigations had been undertaken for the Belgian championship, possibly because of a lack of transparent data. Better transparency will allow constant tracking of clubs' evolution to give them pertinent feedbacks like the report from *Deloitte* (2019) concerning the Economic Impact of the Pro League and further empirical researches regarding the competition.

## 7 Conclusion

Overall, this thesis attempted to determine the determinants for the performance of the clubs engaged in the Belgian professional football championship. An attempt to achieve this, an unbalanced panel data, regrouping 160 observations over ten seasons for twenty-three teams, was built. Thanks to different estimations techniques on the data, we were able to control for potential unobserved team characteristics and endogeneity. In addition, for the sake of backing up the investigations, two proxy variables assessing team performance, for example, league position and point percentage, had been applied. The comparison of the two displayed mainly similar outcomes that impact the seasonal success and should be taken into consideration by football managers. Moreover, two sub-samples regrouping only squads out the G5 and crews who qualified for play-offs one refined the research.

To begin with, we discovered that the causal effect of wage expenditures on Belgian sporting success is positively correlated and verifies the existing literature for alternative championships in Europe (Szymanski Smith, 1997; Simmons Forrest, 2002; Garcia-del-Barrio Szymanski, 2006; Carmichael et al., 2011; Frick, 2013; Llorca Teste, 2016; Ferri et al., 2017; Kuper Szymanski 2018; Madsen et al.,

2018). The effects of other financial variables, such as total incomes and net transfers, are insignificant. Player's salaries are supposed to reflect their true value while purchasing players depends more on the supply and demand from the football market (with no few restrictions in soccer), which can be overvalued (or undervalued) its real value. In addition, because higher team payroll improves team performance, it allows finishing with a higher ranking before the start of the second stage of the competition in Belgium. Consequently, team payroll also has an indirect relationship with success throughout plays-offs 1. Indeed, we have seen that the higher the initial league position when starting them, the better the triumph.

Then, we found counter-intuitive evidence concerning supplementary games during the regular competition. In reality, the findings indicate that when a team reaches the high stage of the Belgian or European cup, they tend to perform better during the season. We attribute this to larger players' motivation that success and create history can involve. Nonetheless, once teams are competing in the play-offs 1, the implication is no more beneficial, and a trade-off comes out for the clubs.

After that, as expected, important attendance during the season is associated with better performance realizations. Clubs need, therefore, to try to maximize their fans' fidelity. However, when we analyzed only the K11 sample, relative revenue appears to be the key determinant to improve team success. Neither team payroll nor extra games nor filling rates are meaningful. We conclude that to win leagues, high salaries are crucial in order to attract the best players, but to remain in the first division, generate revenue is vital to expand its expenses to be able to grow the club.

Afterward, this research also has its own limitations to emphasize. The biggest issue concerns the data used. Unfortunately, they were collected not directly from teams or the league but from external sources. Accordingly, it may be possible that the data set does not correspond at 100% to real figures, especially about wages and incomes, even if we strongly think that they are very close. Bearing this in mind, some estimated coefficients could change. Furthermore, this analysis was specific for the Belgian scheme. Other leagues' estimations could vary.

Finally, this study constitutes one of the first empirical approaches in the Belgian landscape. To strengthen this analysis, similar projects should be achieved. More data accessibility will be a good basis to do so. Further reviews could, for instance, collect data for a larger time period to explore the consequences that implementation of the play-offs system had on competition. We could analyze the factors influencing performance in the old system and make comparisons with the new one. In the same idea, extending data to more divisions in Belgium (at least in the second professional division) makes sense to represent as much as possible the reality.

Finally, this study constitutes one of the first empirical approaches in the

Belgian landscape. To strengthen this analysis, similar projects should be achieved. More data accessibility will be a good basis to do so. Further reviews could, for instance, collecting data for a larger time period to explore the consequences that implementation of the play-offs system had on competition. We could analyze the factors influencing performance in the old system and make comparisons with the new one. In the same idea, extending data to more divisions in Belgium (at least in second professional division) makes sense to represent as much as possible the reality.

Another possible direction for future research is to make use of a non-linear model like the Tobit model since it has the advantage to specifying a range of wanted values for the dependent variable. It is worth noting that in the case of a football championship, there is a minimum number of points, zero, and a maximum number of points depending on the number of games played. Although it has several properties, a Tobit model is likely to be inconsistent with fixed effects. The Poisson model thanks the Quasi-Maximum Likelihood estimator, yields robust results, even in the presence of fixed effects (Wooldridge, 2016).

Furthermore, including a proxy variable to measure the impact of the coach or wage disparity on triumph could be interesting for the *Jupiler Pro League* as various authors did it for other leagues (Franck Nüesch, 2011; Yamamura, 2015; Muehlheusser et al., 2018). Similarly, taking into account the effect of young development can be interesting. An account of a player coming from the academy, who know well the culture of the club and they have nothing to lose, tend to perform well and improve the quality of the team. In addition, they can be considered whether as a good weapon for competition against other clubs.

## Bibliography

- Annala, C., Winfree, J. (2011). *Salary distribution and team performance in Major League Baseball*. *Sport Management Review*, 14, 167–175. <https://doi.org/10.1016/j.smr.2010.08.002>
- Arellano, M., Bond, S. (1991). *Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations*. *The Review of Economic Studies*, 58(2), 277–297. <https://doi.org/10.2307/2297968>
- Bachan, R., Reilly, B., Witt, R. (2014). *Team performance and race: Evidence from the English and French national soccer teams*. *Applied Economics*, 46(13), 1535–1546. <https://doi.org/10.1080/00036846.2013.875108>
- Baltagi, B. H. (Badi H. (2001). *Econometric analysis of panel data (2nd ed.)*. J. Wiley.
- Blundell, R., Bond, S. (1998). *Initial conditions and moment restrictions in dynamic panel data models*. *Journal of Econometrics*, 87, 115–143.
- Buccioli, A., Foss, N. J., Piovesan, M. (2014). *Pay Dispersion and Performance in Teams*. *PLoS ONE*, 9(11), e112631. <https://doi.org/10.1371/journal.pone.0112631>
- Cameron, C., Trivedi, P. (2012). *Microeconometrics Using Stata*. *Economic Record*, 88(283), 595–596. <https://doi.org/10.1111/1475-4932.12006>
- Carmichael, F., McHale, I., Thomas, D. (2011). *Maintaining market position: team performance, revenue and wage expenditure in the English premier league*. *Bulletin of economic research*, 63(4), 464–497. <https://doi.org/10.1111/j.1467-8586.2009.00340.x>
- Caruso, R., Bellavite Pellegrini, C., Di Domizio, M. (2016). *Does diversity in the payroll affect soccer teams' performance? Evidence from the Italian Serie A*. University Library of Munich, Germany. <https://EconPapers.repec.org/RePEc:pra:mprapa:75644>
- Caruso, R., Di Domizio, M., Rossignoli, D. (2017). *Aggregate wages of players and performance in Italian Serie A*. *Economia Politica*, 34(3), 515–531. <https://doi.org/10.1007/s40888-017-0062-6>
- Deloitte. (2019). *Socio-economic impact of the Pro League on the Belgian economy: Determining the footprint of Belgian professional football for the season 17/18 (2nd edition)*. Retrieved from <https://www2.deloitte.com/be/en/pages/technology-media-and-telecommunications/articles/deloitte-pro-league.html>
- European Football Statistics. (2020). *EFS Attendances*. Retrieved November 20, 2019, from <https://www.european-football-statistics.co.uk/attn.htm>
- European Leagues. (2018). *European leagues present its new fan attendance report*. Retrieved May 24, 2020, from <https://europeanleagues.com/epfl-presents-new-fan-attendance-report/>

- Ferri, L., Macchioni, R., Maffei, M., Zampella, A. (2017). *Financial Versus Sports Performance: The Missing Link*. International Journal of Business and Management, 12(3), 36. <https://doi.org/10.5539/ijbm.v12n3p36>
- FIFA. (2018, December 21). *More than half the world watched record-breaking 2018 World Cup*. Retrieved May 5, 2020, from <https://www.fifa.com/worldcup/news/more-than-half-the-world-watched-record-breaking-2018-world-cup>
- Forrest, D., Simmons, R. (2002). *Team Salaries and Playing Success in Sports: A Comparative Perspective (pp. 221–238)*. [https://doi.org/10.1007/978-3-663-07711-4\\_2](https://doi.org/10.1007/978-3-663-07711-4_2)
- Franck, E., Nüesch, S. (2011). *The effect of wage dispersion on team outcome and the way team outcome is produced*. Applied Economics, 43(23), 3037–3049. <https://doi.org/10.1080/00036840903427224>
- Frick, B. (2013). *Team wage bills and sporting performance: Evidence from (major and minor) European football leagues*. The Econometrics of Sport (pp. 63–76). <https://doi.org/10.4337/9781781002865.00012>
- Frick, B., Simmons, R. (2014). *The footballers' labor market after the Bosman ruling (pp. 203–226)*. <https://doi.org/10.4337/9781781003176.00020>
- Garcia-del-Barrio, P., Szymanski, S. (2006). *Goal! Profit maximization and win maximization in football leagues*. AIES-IASE Working Paper Series, 6.
- Gásquez, R., Royuela, V. (2016). *The Determinants of International Football Success: A Panel Data Analysis of the Elo Rating*. Social Science Quarterly, 97(2), 125–141. <https://doi.org/10.1111/ssqu.12262>
- Gautier, G. (2019). *Canicule sur le marché*. Sport Foot Magazine, 19(35), 30–34.
- Hall, S., Szymanski, S., 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. <https://doi.org/10.1177/152700250200300204>
- Hoffmann, R., Ging, L. C., Ramasamy, B. (2002). *The Socio-Economic Determinants of International Soccer Performance*. Journal of Applied Economics, 5(2), 253–272. <https://doi.org/10.1080/15140326.2002.12040579>
- Houston, R. G., Wilson, D. P. (2002). *Income, leisure and proficiency: An economic study of football performance*. Applied Economics Letters, 9(14), 939–943. <https://doi.org/10.1080/13504850210140150>
- Kuper, S., Szymanski, S. (2018). *Soccernomics: Why England loses; why Germany, Spain, and France win; and why one day Japan, Iraq, and the United States will become kings of the world's most popular sport (2018 World Cup Edition)*. Nation Books, an imprint of Perseus Books, LLC.

- Le Soir. (2020). *Droits TV: en cinq ans, Eleven garantit un bonus de 19 millions d'euros à Bruges*. Retrieved April 26, 2020, from <https://plus.lesoir.be/279860/article/2020-02-14/droits-tv-en-cinq-ans-eleven-garantit-un-bonus-de-19-millions-deuros-bruges>
- Llorca, M., Teste, T. (2016). *Dépenses salariales et performance dans l'industrie du football*. *Revue française d'économie*, 21(2), 125. <https://doi.org/10.3917/rfe.162.0125>
- Madsen, D. Ø., Stenheim, T., Boas Hansen, S., Zagheri, A., Grønseth, B. O. (2018). *Wage expenditures and sporting success: An analysis of Norwegian and Swedish football 2010–2013*. *Cogent Social Sciences*, 4(1). <https://doi.org/10.1080/23311886.2018.1457423>
- Marburger, D. R. (1994). *Bargaining power and the structure of salaries in major league Baseball*. *Managerial and Decision Economics*, 15(5), 433–441. <https://doi.org/10.1002/mde.4090150506>
- MLB. (2020). *First-Year Player Draft Rules*. Retrieved May 11, 2020, from <http://mlb.mlb.com/mlb/draftday/rules.jsp>
- Muehlheusser, G., Schneemann, S., Sliwka, D., Wallmeier, N. (2018). *The Contribution of Managers to Organizational Success: Evidence from German Soccer*. *Journal of Sports Economics*, 19(6), 786–819. <https://doi.org/10.1177/1527002516674760>
- NBA. (2020). *NBA Draft Lottery: Schedule, odds and how it works*. Retrieved May 11, 2020, from <https://www.nba.com/nba-draft-lottery-explainer>
- NBB. (2020). *NBB - Central Balance Sheet Office. Online consultation of annual accounts*. Retrieved May 13, 2020, from <https://cri.nbb.be/bc9/web/catalog>
- NFL. (2020). *The Rules of the Draft | NFL Football Operations*. Retrieved May 11, 2020, from <https://operations.nfl.com/the-players/the-nfl-draft/the-rules-of-the-draft/>
- Picazo-Tadeo, A. J., González-Gómez, F. (2010). *Does playing several competitions influence a team's league performance? Evidence from Spanish professional football*. *Central European Journal of Operations Research*, 18(3), 413–432. <https://doi.org/10.1007/s10100-009-0117-z>
- Roodman, D. (2009). *How to do Xtabond2: An Introduction to Difference and System GMM in Stata*. *The Stata Journal*, 9(1), 86–136. <https://doi.org/10.1177/1536867X0900900106>
- Royuela, V., Gásquez, R. (2019). *On the Influence of Foreign Players on the Success of Football Clubs*. *Journal of Sports Economics*, 20(5), 718–741. <https://doi.org/10.1177/1527002518807960>
- Scully, G. W. (1974). *Pay and Performance in Major League Baseball*. *American Economic Association*, 64(6), 15–930.

- Soccerex. (2020). *Belgian Pro League seals 103m domestic rights deal with Eleven*. Retrieved May 27, 2020, from <https://www.soccerex.com/insight/articles/2020/belgian-pro-league-seals-103m-domestic-rights-deal-with-eleven>
- Stiroh, K. (2002). *Playing for Keeps: Pay and Performance in the NBA*. *Economic Inquiry*, 45, 145–161. <https://doi.org/10.2139/ssrn.334421>
- Szymanski, Stefan, Smith, R. (1997). *The English Football Industry: Profit, performance and industrial structure*. *International Review of Applied Economics*, 11(1), 135–153. <https://doi.org/10.1080/02692179700000008>
- Tillou, C., Liarte, S. (2012). *L'impact de l'expérience sur la performance du groupe. Le cas de l'équipe nationale de football du Brésil de 1954 à 2010*. *Management international*, 16(2), 113. <https://doi.org/10.7202/1008711ar>
- Transfermarkt. (2020). *Jupiler Pro League*. Retrieved March 22, 2020, from <https://www.transfermarkt.com/jupiler-pro-league/startseite/wettbewerb/BE1>
- UEFA. (2010-2020). *The European Club Footballing Landscape - Financial Year 2009-2019*. (3-11). Retrieved from [https://www.uefa.com/MultimediaFiles/Download/OfficialDocument/uefaorg/Clublicensing/02/63/79/75/2637975\\_DOWNLOAD.pdf](https://www.uefa.com/MultimediaFiles/Download/OfficialDocument/uefaorg/Clublicensing/02/63/79/75/2637975_DOWNLOAD.pdf), pp. 56-81.
- UEFA. (2019a). *Financial Fair Play | Inside UEFA*. Retrieved April 29, 2020, from <https://www.uefa.com/insideuefa/protecting-the-game/financial-fair-play/>
- UEFA. (2019b). *UEFA Champions League 2018/19: payments to participating clubs*. Retrieved April 29, 2020, from <https://www.uefa.com/uefachampionsleague/news/025a-0e9f90234d35-f3f52c29a815-1000-uefa-champions-league-2018-19-payments-to-participating-clubs/>
- UEFA. (2019c). *UEFA Europa League 2018/19: payments to participating clubs*. Retrieved April 29, 2020, from <https://www.uefa.com/uefaeuropaleague/news/025a-0e9f90158751-7513a60fc3f2-1000-uefa-europa-league-2018-19-payments-to-participating-clubs/>
- Wooldridge, J. M. (2016). *Introductory econometrics: A modern approach (Sixth edition)*. Cengage Learning.
- Yamamura, E. (2015). *Wage Disparity and Team Performance in the Process of Industry Development: Evidence From Japan's Professional Football League*. *Journal of Sports Economics*, 16(2), 214–223. <https://doi.org/10.1177/1527002512471539>

## Appendix

### **A.1 Clubs who took part in the *Jupiler Pro League* since 2009-2010 until 2018-2019**

Anderlecht (10), Antwerp (2), Beerschoot (4), Cercle Brugge (7), Charleroi (9), Club Brugge (10), Eupen (4), Genk (10), Gent (10), Kortrijk (10), Lierse (5), Lokeren (10), Mechelen (9), Mons (3), Mouscron (6), OHL (3), Oostende (6), Roeselare (1), Sint Truiden (7), Standard de Liège (10), Waasland Beveren (7), Westerlo (6) and Zulte Waregem (10)<sup>22</sup>.

### **A.2 Clubs who played the play-offs 1 since their introduction until 2018-2019**

Anderlecht (10), Antwerp (1), Charleroi (3), Club Brugge (10), Genk (7), Gent (8), Kortrijk (3), Lokeren (3), Oostende (2), Sint Truiden (1), Standard Liege (7) and Zulte Waregem (5)<sup>23</sup>.

---

<sup>22</sup>The number in parenthesis shows how many times the club participated in the JPL since 09-10 until 18-19

<sup>23</sup>The number in parenthesis indicates how many times the club qualified for plays-offs 1.

### A.3 Financial evolution of the *Jupiler Pro League*

Table 7: Average financial evolution by season (in millions of euros)

season	payroll	net trans	tv rev	comm rev	euro rev	gates rev	total rev
09-10	8.515	0.204	2.628	2.552	1.206	13.699	20.085
10-11	10.510	-1.469	2.732	2.684	0.336	14.336	20.088
11-12	11.349	-4.008	3.209	2.453	1.177	12.710	19.549
12-13	11.271	-0.537	3.096	2.793	1.225	9.705	16.819
13-14	11.973	-0.949	3.033	3.115	1.206	10.351	17.705
14-15	14.073	-2.088	3.859	4.512	1.759	10.704	20.834
15-16	15.273	-3.604	3.885	5.359	2.598	11.104	22.945
16-17	16.358	-2.413	3.818	6.116	2.602	10.759	23.296
17-18	15.913	-4.684	4.371	6.351	1.447	10.973	23.142
18-19	17.255	-1.928	4.254	6.223	3.186	10.185	23.847
Total	13.249	-2.148	3.488	4.216	1.674	11.453	20.831

Source: Own data analysis drawn from Stata 14.

Globally, we notice that there is an upward trend in every section unless for European and gates revenue where the tendency is quite stable.

The biggest augmentation appears in wages. Indeed, since 2009, they have increased by an impressive 102.64% (in real prices). The largest jump occurred between season 2013-2014 and the beginning of the season 2014-2015 (+ 17%). This phenomenon can be explained by the fact their clubs signed a new TV-broadcasting in 2014-2014 and commercial experienced its major growth (+44%).

Then, regarding the variation of transfers, table 7 confirms that the *Jupiler Pro League* transfers is an important source of revenue for clubs. Since season 2010-2011, the average net transfers has always been negative.

## A.4 Correlation matrices complete sample

Table 8: Correlation matrix with league position

Var	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
league_pos (1)	1.000									
rel_wages (2)	0.702	1.000								
rel_rev (3)	0.746	0.837	1.000							
trans_net (4)	-0.181	-0.237	-0.199	1.000						
av_age (5)	0.041	-0.009	-0.058	0.045	1.000					
foreigners_prop (6)	0.219	0.299	0.247	-0.078	0.300	1.000				
cup_games (7)	0.277	0.139	0.186	-0.049	0.125	0.026	1.000			
euro_games (8)	0.644	0.705	0.801	-0.296	-0.103	0.134	0.140	1.000		
roster (9)	0.021	0.231	0.279	-0.118	-0.155	0.318	-0.125	0.269	1.000	
filling_rate (10)	0.591	0.496	0.647	-0.107	0.225	0.151	0.193	0.460	0.122	1.000

Source: Own data analysis drawn from Stata 14.

Table 9: Correlation matrix with point percentage

Var	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
pnt_pct (1)	1.000									
rel_wages (2)	0.656	1.000								
rel_rev (3)	0.715	0.837	1.000							
trans_net (4)	-0.190	-0.237	-0.199	1.000						
av_age (5)	0.152	-0.009	-0.058	0.045	1.000					
foreigners_prop (6)	0.270	0.299	0.247	-0.078	0.300	1.000				
cup_games (7)	0.275	0.139	0.186	-0.049	0.125	0.026	1.000			
euro_games (8)	0.597	0.705	0.801	-0.296	-0.103	0.134	0.140	1.000		
roster (9)	0.006	0.231	0.279	-0.118	-0.155	0.318	-0.125	0.269	1.000	
filling_rate (10)	0.618	0.496	0.647	-0.107	0.225	0.151	0.193	0.460	0.122	1.000

Source: Own data analysis drawn from Stata 14.

## A.5 Relation team payroll on team performance when using point percentage as DV

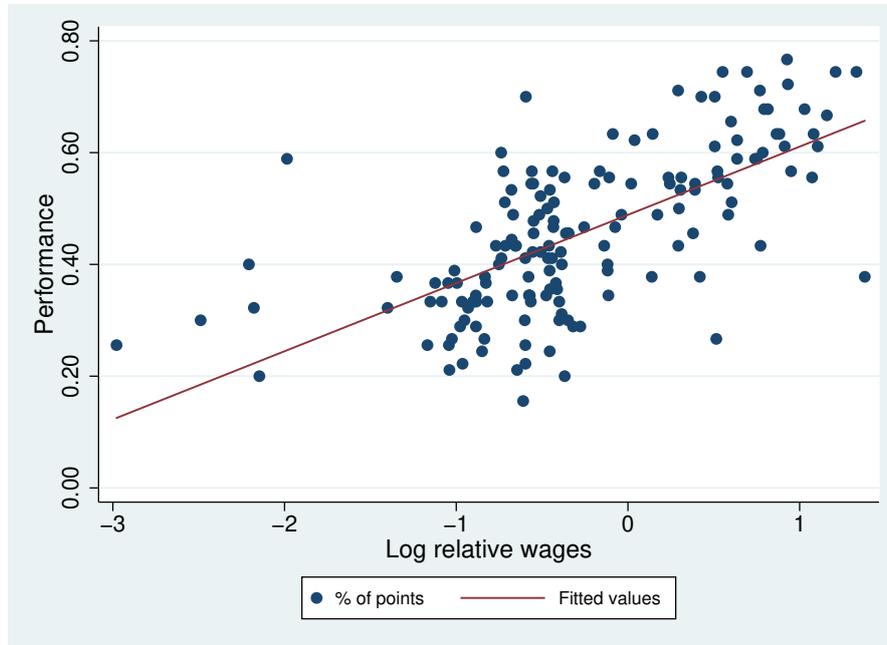


Figure 3: Relation team payroll on team performance (2).

Source: Own data analysis drawn from Stata 14.

## A.6 Additional regressions

Table 10: Decomposition of the main regression by pooled OLS (1)

Dependent variable: league position	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
lrel_wages	<b>0.420***</b> (0.067)	<b>0.140**</b> (0.050)	<b>0.139**</b> (0.050)	<b>0.132***</b> (0.043)	<b>0.134***</b> (0.047)	<b>0.128***</b> (0.039)	<b>0.110***</b> (0.039)	<b>0.124***</b> (0.040)
lrel_rev		<b>0.477***</b> (0.086)	<b>0.477***</b> (0.086)	<b>0.521***</b> (0.057)	<b>0.520***</b> (0.062)	<b>0.424***</b> (0.091)	<b>0.450***</b> (0.084)	<b>0.348***</b> (0.100)
trans_net			-0.000 (0.003)	-0.001 (0.003)	-0.001 (0.003)	0.000 (0.003)	0.000 (0.003)	0.000 (0.003)
av_age				<b>2.323**</b> (0.930)	<b>2.311**</b> (0.894)	<b>2.322**</b> (0.886)	<b>1.875**</b> (0.810)	<b>1.727*</b> (0.846)
sqav_age				<b>-0.048**</b> (0.020)	<b>-0.047**</b> (0.019)	<b>-0.048**</b> (0.019)	<b>-0.039**</b> (0.017)	<b>-0.036*</b> (0.018)
foreigners_prop					-0.017 (0.192)	0.047 (0.212)	0.270 (0.210)	0.298 (0.203)
cup_games						<b>0.027*</b> (0.013)	0.020 (0.013)	<b>0.020*</b> (0.011)
euro_games						0.014 (0.010)	0.016 (0.009)	<b>0.018*</b> (0.009)
roster							-0.026 (0.031)	-0.020 (0.030)
sqroster							0.000 (0.000)	0.000 (0.000)
filling_rate								<b>0.279*</b> (0.158)
_cons	<b>0.425***</b> (0.049)	<b>0.423***</b> (0.041)	<b>0.422***</b> (0.042)	<b>-27.769**</b> (10.866)	<b>-27.629**</b> (10.420)	<b>-27.886**</b> (10.331)	<b>-21.589**</b> (9.553)	<b>-20.011*</b> (9.961)
N	159	159	159	159	159	159	159	159
adj. $R^2$	0.472	0.599	0.596	0.625	0.623	0.635	0.654	0.657
Season FE	no	no	no	no	no	no	no	no
Team FE	no	no	no	no	no	no	no	no

Clustered robust standard errors in parentheses

\* p<0.1, \*\* p<0.05, \*\*\* p<0.01

Source: Own data analysis drawn from Stata 14.

We notice that every time a new explanatory variable is included, the adjusted  $R^2$  rises. This is not surprising and underlines the importance of controlling other relevant factors. However, only the variables relative wage and relative revenue account for 60% of the variation in league position. Financial variables are clearly deterministic.

Table 11: Decomposition of the main regression by pooled OLS (2)

Dependent variable: point percentage	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
lrel_wages	<b>0.272***</b> (0.038)	<b>0.093**</b> (0.044)	<b>0.091*</b> (0.045)	<b>0.089**</b> (0.039)	<b>0.081*</b> (0.042)	<b>0.079*</b> (0.041)	0.060 (0.039)	<b>0.080**</b> (0.038)
lrel_rev		<b>0.306***</b> (0.062)	<b>0.304***</b> (0.062)	<b>0.328***</b> (0.051)	<b>0.332***</b> (0.055)	<b>0.317***</b> (0.063)	<b>0.344***</b> (0.059)	<b>0.201**</b> (0.075)
trans_net			-0.001 (0.002)	-0.002 (0.001)	-0.002 (0.001)	-0.001 (0.001)	-0.002 (0.002)	-0.002 (0.002)
av_age				<b>1.627*</b> (0.812)	<b>1.687**</b> (0.793)	<b>1.694**</b> (0.801)	<b>1.267*</b> (0.620)	1.059 (0.637)
sqav_age				<b>-0.034*</b> (0.017)	<b>-0.035*</b> (0.017)	<b>-0.035*</b> (0.017)	<b>-0.027*</b> (0.013)	-0.023 (0.014)
foreigners_prop					0.092 (0.179)	0.117 (0.181)	<b>0.337*</b> (0.189)	<b>0.376**</b> (0.172)
cup_games						0.017 (0.013)	0.011 (0.012)	0.011 (0.010)
euro_games						0.001 (0.005)	0.002 (0.005)	0.006 (0.005)
roster							-0.030 (0.025)	-0.022 (0.022)
sqroster							0.000 (0.000)	0.000 (0.000)
filling_rate								<b>0.392***</b> (0.108)
_cons	<b>-0.764***</b> (0.028)	<b>-0.765***</b> (0.023)	<b>-0.768***</b> (0.025)	<b>-20.412**</b> (9.520)	<b>-21.167**</b> (9.291)	<b>-21.290**</b> (9.385)	<b>-15.167**</b> (7.056)	<b>-12.951*</b> (7.259)
<i>N</i>	159	159	159	159	159	159	159	159
adj. <i>R</i> <sup>2</sup>	0.399	0.503	0.500	0.522	0.520	0.523	0.562	0.583
Season FE	no	no						
Team FE	no	no						

Clustered robust standard errors in parentheses

\* p<0.1, \*\* p<0.05, \*\*\* p<0.01

Source: Own data analysis drawn from Stata 14.