Owning the Mona Lisa: The influence of changes in the rules of sport organisations on the value of players

Dissertation zur Erlangung des Doktorgrades an der Fakultät Rechtswissenschaft der Universität Hamburg

vorgelegt von
Jean-Christian Drolet
aus
Roberval

Hamburg 2011
To Annika
Table of content

Acknowledgements ........................................................................................................................................ VI

Introduction ................................................................................................................................................ 1
   I. The goal of this thesis .............................................................................................................................. 2
   II. The method ........................................................................................................................................... 3

Chapter 1: The evolution of professional sport organisations ................................................................. 7
   1.1. The economic structure of professional sports ............................................................................... 8
       1.1.1. The product of sport .................................................................................................................. 8
           1.1.1.1. The game ............................................................................................................................ 8
           1.1.1.2. The championship or beyond the games ........................................................................... 10
       1.2. What do leagues actually do? ....................................................................................................... 12
           1.2.1. The “team goods” of sport ..................................................................................................... 12
       1.3. The economic nature of sport leagues ......................................................................................... 14
           1.3.1. Sport as a multi-plant firm in a Natural Monopoly market ..................................................... 15
           1.3.2. The league as a joint venture .................................................................................................. 16
               1.3.2.1. The single entity doctrine ................................................................................................. 17
               1.3.2.2. The single entity doctrine in sports .................................................................................. 18
               1.3.2.3. The single entity doctrine in the EU .............................................................................. 19
               1.3.2.4. The end of the joint venture ............................................................................................ 19
           1.3.3. The cartel .................................................................................................................................. 20
               1.3.3.1. The consequences of cartelisation of sports ...................................................................... 21
               1.3.3.2. Why not break up the cartel and let the market do the job? ........................................... 22
       1.4. The genesis of professional sport leagues ...................................................................................... 24
           1.4.1. Barnstorming ............................................................................................................................ 24
           1.4.2. The tournaments ...................................................................................................................... 24
           1.4.3. The Leagues ............................................................................................................................. 25
               1.4.3.1. How professional leagues started ...................................................................................... 25
               1.4.3.2. England and the case of soccer ............................................................................................ 26
               1.4.3.3. Major League Baseball .................................................................................................. 29
                   1.4.3.3.1. The birth of the National League ................................................................................. 30
                   1.4.3.3.2. Closed leagues ............................................................................................................ 31
Chapter 2: Competitive Balance or Uncertainty of result

2.1. The need for uncertainty

2.2. The Louis-Schmeling paradox
   2.2.1. Holes in the paradox

2.3. What is uncertainty of result in a professional sport league context?
   2.3.1. Game competitive balance
   2.3.2. Championship competitive balance
      2.3.2.1. Single season standard deviation ratio
      2.3.2.2. Championship variance balance
   2.3.3. The Blue Ribbon Panel on Baseball Economics definition

2.4. Competitive Balance and its effect on players

Chapter 3: The determination of value in professional team sports

3.1. The situation as it stands now for determining the value

3.2. The value of teams
   3.2.1. The market value of teams
   3.2.2. The publicly traded corporation in sports
   3.2.3. US teams, publicly traded corporation and their value

3.3. The sportsman owner
   3.3.1. The leveraged syndicate
   3.3.2. The sportsman owner and the leveraged syndicate
   3.3.3. The empirical problem with the sportsman owner
   3.3.4. European soccer and the sportsman owner

3.4. What the value of sports teams really is about

3.5. Determining the value of professional sport players
   3.5.1. Determining the player’s value with the revenue function of the team
   3.5.2. The willingness to pay for the services of a player
   3.5.3. The surplus of teams and its influence on the willingness to pay for players
   3.5.4. Multiple years contracts

3.6. The definition of value of a player for the purpose of this work

Chapter 4: The effect of a change in the rules of the game on the value of players

4.1. The rules of the game
   4.1.1. Definition of the rules of the game
5.3.2. Contractual uncertainty ........................................................................................................ 94
5.3.3. The negotiated reserve system ............................................................................................ 94
  5.3.3.1. Rookie non-free agents .................................................................................................. 94
  5.3.3.2. Restricted free agents .................................................................................................. 95
    5.3.3.2.1. The right of first refusal ..................................................................................... 96
    5.3.3.2.2. Compensation ...................................................................................................... 96
    5.3.3.2.3. The effects of restricted free agency on the value of players ......................... 96
  5.3.3.3. Unrestricted free agents ............................................................................................ 98
  5.3.3.4. Special cases: the franchise and transition player designation ......................... 98
5.4. Entering the players pool, the college draft ......................................................................... 99
  5.4.1. Training and development of young players in football .............................................. 99
  5.4.2. The NFL college draft ..................................................................................................... 100
  5.4.3. Contesting the draft ....................................................................................................... 101
  5.4.4. Empirical evidence regarding the influence of the draft on competitive balance in the NFL ................................................. 102
  5.4.5. The draft, a winner’s curse? ......................................................................................... 105
  5.4.6. What is the draft really about? .................................................................................... 107
5.5. The salary cap ...................................................................................................................... 107
5.6. Under the shadow of the law ............................................................................................ 108
5.7. Effect of these measures on the value of players .............................................................. 109
  5.7.1. The impact of a CBA on the value of a player .............................................................. 109
  5.7.2. The impact of free agency on the value of players ...................................................... 110
  5.7.3. The impact of salary caps on the value of players ...................................................... 111
5.8. The impact of a change in the commercial rules of the game in football on the value of players .............................................................................................................................. 112
5.9. Soccer ............................................................................................................................. 113
  5.9.1. The Bosman case ......................................................................................................... 114
    5.9.1.1. History of the case .................................................................................................. 114
    5.9.1.2. What are transfer fees .......................................................................................... 115
    5.9.1.3. The pre-Bosman transfer system ......................................................................... 115
    5.9.1.4. After Bosman ....................................................................................................... 117
  5.9.2. The birth of the new system ......................................................................................... 118
  5.9.3. The Bosman test .......................................................................................................... 118
  5.9.4. The Monti rules ......................................................................................................... 120
5.9.4.1. Analysis of the Monti rules .............................................................. 123
5.9.5. The 2005 transfer rules ........................................................................ 125
  5.9.5.1. The changes .................................................................................. 125
  5.9.5.2. Special changes for Europe .......................................................... 127
5.10. Effect of the transfer rules on the value of soccer players .................... 128
  5.10.1. The direct influence of the new transfer system on the value of the players . 133
  5.10.2. The indirect influence of the rules on the value of the players via competitive balance ................................................................. 134
  5.10.3. The instability of the new system .................................................... 136
5.11. Conclusion: The impact of the change in the rules of the commercial game on the value of players ........................................................................ 142
Conclusion .................................................................................................. 144
Final thoughts .............................................................................................. 150
Bibliography ................................................................................................. 152
Acknowledgements

Although only my name figures here as author, I am indebted to many persons that helped me along the way and made that work possible. This is the moment to thank them.

First and foremost my wife Annika, to whom that work is dedicated, as she also had to endure the ups and downs and creative crisis that came with writing that work. Additionally she was the one that provided me with the first straight face test for my work. I could never have done it without her.

To my in-laws, Thomas and Marlene, for allowing me to be such a parasite for numerous years and never complaining about it.

I must also thank Prof. Dr. Thomas Eger who acted as supervisor to that work and helped me considerably in creating the structure of that work turning it into something more than a collection of eclectic texts and Prof. Dr. Roland Kirstein who provided a very enlightening analysis of my work during our thesis seminar.

The staff, students and the Director of the Doctoral College of Law and Economics of the Universität Hamburg, Prof. Dr. Hans-Bern Schäfer, are also part of that work since their financial, technical and moral support allowed me to concentrate myself fully on my work.

A special thanks to Brian, John, Jack, Klaus, Steve, David and Rob, established scholars and members of the IASL, who tolerated my presence among them and gave me the courage to pursue my work by pointing the novelty of the subject.

Of the people that generously accepted to answer my questions, a special mention here must be made of Mr. Julien Brisebois, vice-president to hockey operations of the Montreal Canadien and now assistant general manager with the Tampa Bay Lightning, who took a whole morning to answer my sometimes naïve questions. His honesty and willingness to share information was priceless.

Last but not least I would like to thank all my family for their support, but most notably:

Danny, for taking the time to speak to me despite his busy schedule, for giving me a place to work in Canada and for all the nice things he said about me and my work; I am starting to believe it myself.

My Mom, Joanne, for always worrying about me and always giving me her unconditional love.

And finally my Dad, Bernard, for always listening to my ideas, helping me to find my own way around problems, for keeping me posted on various sporting events back home and for having postponed indefinitely his ride with La Fille en Bicycle.

However, any mistakes remain my own.
Introduction

"Now that no one buys our votes, the public has long since cast off its cares; the people that once bestowed commands, consulships, legions and all else, now meddles no more and longs eagerly for just two things----Bread and Games!"

-Juvenal, Satire 10

As this well known quote from the Roman satirist shows us, organised spectator sports consisting of athletes competing against one another seem to have attracted large crowds, stoked the interest of the masses and provided numerous business opportunities in Roman society. All human organised societies, from the Aztecs to Victorian England, seem to have used spectator sport for reasons of their own. The consequence of this use of sport was the creation of organisational structures designed by these societies to achieve their goals.

With such organisations came the need for rules created to regulate them, the game and everything else surrounding the sport. One of the first and most important steps was to make sure that everyone was actually playing the same game.

With this emphasis on rules, it was a natural thing for jurists and lawyers to take an interest in sports. With the jurists came a refinement of these rules and regulations. Over the years it became impossible to imagine sport without lawyers. They are everywhere; as agents, managers, even executives and owners. As a consequence, organised sports now have a very sophisticated legal environment with rules getting more and more complicated every year. This results in the creation of a complex alternative legal system that may not be justified by the economic importance of the industry. Even using the most generous estimates, the entire sport industry does not represent more than 1, 5% of the GDP of the United States. If Juvenal was writing his satires today, he might add one about lawyers being interested in sports in order to be allowed to speak about it on the job.

---

I. The goal of this thesis

Professional sport organisations have argued for a long time that they have a special status in society and therefore should be treated differently from other commercial ventures. In the United States, the debate was started by the Supreme Court when it agreed to exempt baseball from anti-trust law\(^2\). The result of this decision gave a commercial advantage to a baseball league and therefore all US based leagues also tried to benefit from that decision. Eventually they all failed, were denied a special status and treated like any other business. In Europe, especially at the EU level, the debate also seemed to have been closed by the European Court of Justice (ECJ) with the *Bosman* and *Mecca Medina* cases\(^3\) that refused to recognise the “special” status of sport. It now seems to have moved to the political arena. The leading professional sport organisations in Europe, led by FIFA, lobbied intensively in order to enshrine what they call the “specificity of sport” in EU legislation\(^4\). They are now fighting to force their interpretation of this legislation through the various EU entities. Their understanding is that this “specificity of sport” should inoculate sport organisations against the unwanted and disturbing effects of “normal” law on their way of doing business.

It is therefore appropriate to ask if sport is indeed special and if this specificity warrants a special legal treatment.

This thesis tries to answer part of that question by trying to explore a limited part of this issue and to determine if there is a specificity of sport in that limited field that would require a special legal treatment. The limited scope of this thesis will be to look at the effect of the regulation of sport organisations on professional sportsmen. The method used will be to approach this problem from a value added perspective, by asking if the legal intervention actually created or destroyed value for the players. This would provide a better measure of the impact of the rules than the declarations of the sports authorities that always affirm that changes are made for the good of the sport. By looking at the true impact of the rules on players, it should allow outside observers to see if the professional sport organisations need supervision or if they can be left to their own device in order to have a positive impact on society as this is basically what they are asking with the “special” status: to be left alone and


\(^3\) *Bosman* will be extensively analysed in Chapter 5 and *Mecca Medina* in Chapter 4.

\(^4\) Treaty of Lisbon amending the Treaty on European Union and the Treaty establishing the European Community, signed at Lisbon, 13 December 2007, article 2, line 123 to 125.
allowed to run their business as they see fit. With that exercise it will allow us to lift part of
the veil and discover if the “specificity of sport” really exists or if it is only a way to try to
lobby for a more favourable legal environment for sport organisations.

Why should the players and not another set of data be examined? This choice can be
explained by a number of factors. The players represent the main assets of the industry,
excluding the stadiums. They also are the biggest recurring expense/investment for the teams.
Players and their importance are universal. In all sports they have to contribute to the sporting
and commercial success of their respective teams. Without the players there is no professional
sport business. The data set “players” therefore allows for a comparison of the different sports
and organisations, an exercise that is usually difficult because of the diversity of the industry.
With the players we find a common denominator that is spread across sports. In addition,
information about the performances and revenues of players is easier to obtain than
information about any other participant of the sport industry. Their salaries and performance
are well publicised, observable and the subject of numerous statistics and analyses.

From a legal perspective, the relationship between players, teams, value and the law is one of
the best documented in the industry. Legal conflicts in sport were almost exclusively the
result of a breakdown of the relationship between the players and the teams. These numerous
conflicts resulted in changes that were imposed on the sport industry. This allowed observers
to determine the impact of these changes on various aspects of the industry.

In this thesis I will examine these changes and try to determine how they affected the value of
players. I will try to determine if there is some “sport social welfare” or improvement of the
industry as a whole, coming from the changes in rules. The second step will consist in
determining if the change is “special”, that means: could this change be considered illegal
using the normal legislation. By identifying the beneficiary of the change and the impact on
the industry it will be possible to see if letting the sport industry regulate itself is a good idea.

II. The method

In this thesis I will examine these changes and try to determine how they affected the value of
players. I will try to determine if there is some “sport social welfare” or improvement of the
industry as a whole, coming from the changes in rules. The second step will consist in
determining if the change is “special”, that means: could this change be considered illegal
using the normal legislation. By identifying the beneficiary of the change and the impact on
the industry it will be possible to see if letting the sport industry regulate itself is a good idea.

II. The method

In chapter 1 I will examine the evolution of the professional sport organisations. The goal of
this chapter is to give the reader a general review of the current economic and legal theories
on professional sports and to try to discover if there are unique characteristics to that industry.
In addition, this chapter will provide a short history of professional sport leagues where I will examine the parallels and divergences leagues developed over the years. At the end of this chapter, the reader should have an understanding of the economics and legal situation of the sport industry.

Chapter 2 will examine one of the most important aspects of the business of sport: uncertainty of result, also called competitive balance. According to traditional sport economic theory, uncertainty creates excitement for the fans and therefore a higher demand for the sport. The harder it is to predict the result of a game, the greater the number of fans who want to see it. The purpose of this chapter is to present a synthesis of the most unique and important aspects of the business of sport, since the majority of the regulations in sport are justified by the authorities in order to promote the competitive balance. I will start by looking at what is called the Lewis-Schmeling paradox and how we can apply it to the context of this thesis. After that, I will look more closely at competitive balance in the context of a professional sport league and some of the methods that have been proposed to define it. In the end I will settle on a definition for the purpose of this thesis that is less about the right calculation and more about helping us to see the impact of it on the business of professional sport leagues.

Chapter 3 will examine the notion of value in professional team sports. The main problem when trying to determine the value of teams in professional sports is the lack of reliable information and the culture of secrecy that surrounds the industry. The best an outsider can achieve is an estimate. The value of teams will be analysed in order to understand how the industry sees the concept of value. With that sport understanding of the concept of value, I will try to apply it to the players. Determining the value of players seems easy, but this is deceptive. The salaries of the players and the transfer fees paid for them are usually made public, either through the collective bargaining agreement or with leaks to the press, but does that really represent the value of the players? Finding an objective way of measuring their value is hard, especially in an industry where the performances of players are measured with traditional statistics but also with the subjective opinion of managers, coaches and agents. What if all these things are based on false assumptions? Is there actually a way to take into consideration the subjective judgements that seem to be required to measure the performance of players? Also, are there other factors than the on-field performances that help determine the value of a player? In the commercial era that professional sports are now going through, the pure commercial value of a player should be at least considered. At the end of the chapter, I
will propose a definition of value that will take all these factors into consideration, remain relatively simple and be easy to understand in order to help the reader comprehend the arguments that are going to be presented in the following chapters.

Chapter 4 will examine the effect of a change in the rules of the game on the value of players. Rules of the game means the rules by which the sport is played on the field. I will first examine the reasons that motivate sport organisations to change the way they play. Since these changes are rare it is important to understand what motivates the sport authorities to make them. I will show that most of changes in the rules of the game are implemented for two main reasons: to increase the safety of the participant and/or to generate more interest for the spectators. I will also look at the types of rule changes that can happen, since not all rules changes have the same impact. I will show that there are two types of change in the rules that can happen, the radical and the mild rule change. The main focus will be to determine, for both types of changes, what is the consequence on the value of players. Are all players affected equally? If there is a growth in revenues are the players going to get their share? To try to answer these questions I will use and analyse some concrete example of radical and mild rule change. These examples will be taken from North American sports since this is where the best documented cases of change in the rules of the game are to be found. In the end of this chapter, I should be able to see what kind of impact both types of changes have on the value of the players.

In chapter 5, I will examine how the changes in the rules of the commercial game affect the value of the players. The rules of the commercial game are the internal rules of the sport, not related to the game on the field, that are used to regulate the relationship between the different parties of the sport business. These rules regulate the relationship between the league and the teams, between the teams and between the teams and the players. These have been the most important influence on the value of the players and the subject of most legal conflicts in the industry. I will mainly look at two sports: soccer and football and see if the different ways in which they deal with the players in the commercial rules of the game have the same or a different impact on their value. Football is using the traditional North American model of collective bargaining agreement since it faces players represented by a strong and active union. The commercial rules of the game of football have therefore been determined through

---

5 Being Canadian, in this thesis I will use in this thesis the term most familiar to me, soccer, for what the rest of the world knows as football. The term football will stand for the sport known to the rest of the world as American football.
negotiation. For soccer, the commercial rules of the game have been imposed in a top down fashion by the central authority of the sport, Fédération Internationale de Football Association, FIFA. This comparison will allow us to see if there is a different effect on the value of the player coming from rules that were negotiated or from rules that were imposed. At the end of this chapter it will be pretty clear that the impact of a change in the commercial rules, independent of the way it is implemented, is likely to be a redistributive one and that growth rarely originated from it. However, the way the rule was implemented will have an impact on the way the wealth is redistributed between the players.

In conclusion, I will show that the changes in rules examined in that thesis result in a very specific type of value change and I will discuss how further research can orient itself in order to build on the results found in this thesis.
Chapter 1 – The evolution of professional sport organisations

If all the year were playing holidays,
to sport would be as tedious as work;

- William Shakespeare, Henry IV, Part I

Sport may be the oldest form of distraction created by mankind. Most sports asked the participant to repeat the movements or tasks that they had to do in everyday life. The difference with sport is that the results *per se* are measured, not the products coming from the feat itself. It is not the fact that the javelin hits an animal or an enemy that matters, but the distance at which it was thrown or the precision with which one hits a target. The ancient Greek and Roman games, the ancestors of our modern sports, were based on skills needed during war. Even in the modern era that spirit continued with modern Pentathlon that is basically a rehearsal of the skills needed by the XIX\(^{th}\) century military officers\(^6\). Some even go further saying that international sport competitions serve as an ersatz for war\(^7\).

The exact moment when professional sports -sports in which the participants are paid to play the game- made their appearance is still a matter for debate. It is hard to determine the first time someone was paid to play a sport. But it is generally agreed that the professional sport organisations that we know today were created at the end of the XIX\(^{th}\) century\(^8\). This is the moment in time where leagues, teams and sport organisations were shaped and started to resemble what we know today.

In this first chapter will first examine the economic and legal nature of professional sport. This analysis will allow me to make preliminary observations as to the nature of this business. The goal is to see if there is a difference between the business of sport and other ordinary commercial ventures. Second, I will examine the historical evolution of sport leagues from their first organisational forms to the actual modern leagues. Finally, I will try to define what leagues are in economic terms to help qualify the characteristics of these organisations.


\(^8\) More on the start of professionalism later in this chapter.
goal of this first chapter is to provide the reader with a basic understanding of the economics and history of professional sport in order to have a clear perception of the organisations and structures that create the rules and hire the players.

1.1. The economic structure of professional sports

It seems that each time there is a lawsuit implicating a sports club or league, their first defence is to invoke an exemption to the rule of law because of some “special” status of sport. If the law is applied normally, they argue, that would result in the destruction of professional sport as we know it. In some cases they were successful; in other cases they had, like any other industry, to obey the law. However, it is a legitimate question to ask if there is a fundamental difference between the business models of professional sport and the other “normal” business. Like it was so clearly put to me by Stephen Weatherill: “Is there any fundamental difference between operating a sausage stand and a soccer team?” In order to at least provide a partial answer in the context of this work, I will first look at the economic nature of the professional sport industry, starting with its product.

1.1.1. The product of sport

“What does sport produce?” This basic question is at the origin of the claim that the industry is special. Unlike other industries, sport does not produce goods; one could even argue that there are no services produced either. This difficulty to find the product of sport has been used by the partisan of the specificity of sport to justify their position. However, in the case of professional team sports, there is at least one clear product that we can identify: the game.

1.1.1.1. The game

The game is what results when two teams meet to face each other in an athletic competition. This is the basic unit of the professional sport team business model. Without the game, there would be nothing to watch in the stadium, no merchandising, no television or radio broadcasting contract, etc. It is sport business at its simplest expression, just one team against

---

10 Federal Baseball Club of Baltimore, see note 2.
11 Radovich and Bosman, see note 9.
12 Interview with the Stephen Weatherill, Professor, Faculty of Law Oxford University, November 2007.
another for one match to determine who the best is. Without the game, there are no professional sports.

Usually, a business is the sole manufacturer of the product that it puts on the market. It controls all the inputs that lead to the end product and does not need help in order to put its product out in the market. All the other firms producing the same product are competitors, fighting each other for a share of the market. There can be some kind of interactions between competing firms, joined ventures or collaboration with them, but in the end, it is the single firm that is the sole entity responsible for the creation of the end product. The situation is different when it is time to organise a game.

In order to set up a game, the team needs an opponent, another team to which it can measure itself. The foundation of the professional sport model is based on the idea that nobody will come to see a team playing alone. A team that stands alone will disappear unless it can find opponents to compete against. These opponents also represent competitors since both teams produce the same output and fight for the same consumers. But they need each other in order to create their product since one team represents only one of the two sides of the game.

If trying to present the production function of a professional sport game in a very simple mathematical equation, one could do it this way:

$$PT1 \times PT2 = GP$$

$PT1$: Production of team 1
$PT2$: Production team 2
$GP$: Game production

If one team is absent from the field, i.e. $PT=0$, then there is no game and in consequence the value of $GP$ will be equal to zero. By combining their production factors, both teams create the game, “the product” which is consumed in professional sport. This industry may be the only one in which the firms need the absolute collaboration of their competitors to make a profit. This fact also has an impact on the profit maximising model for the professional sport teams. For the traditional firms, the model that allows them to maximize their profit is the monopoly. A monopolist, being the only one who produces a good, is capable of producing
the level of goods or imposing a price that allows him to maximise his revenues at the
detriment of social welfare. A monopoly will always result in a suboptimal situation for the
society as a whole\textsuperscript{13}.

In the case of a professional sport team, as shown by the previous production function, being
a monopolist in the production of games cannot be profit maximising. It is in fact the worst
possible situation for a team since it results in a model with no revenues: no opponents means
no revenues. The presentation of games in the professional sport industry is one of the only
economic situations in which a monopoly is an unviable model. So in a sense, there is
specificity in sport as the nature of the market for games requires a strong interaction between
competitors. But the games are not the only product in professional sport. There is a second
product that professional sport can exploit: the championship.

1.1.1.2. The championship or beyond the games

As the market for professional sports grows and teams multiply, it becomes difficult to
determine which one is the best team. How do you determine the official “ranking of teams”
so that the team of X or Y can boast having the best team in the country? This cannot be
determined by individual games as the multiple individual match-up possibilities make it hard
to determine who the best is. This situation creates the favourable conditions for the
emergence of a new market, the market for championship. Championship should be
understood as a way of determining the best team in a sport in a pre-determined area.

In a championship, each game has an added value that is a consequence of it being played in
that context. The game, taken individually, has its own value derived from the quality of the
teams on the field, the abilities of the players, etc. In a championship context, one may add to
that the impact of the game on the ranking of the championship. The result of the game will
affect the standings of the implicated teams in the championship and may have an impact on
the other teams in the league that are not implicated in the game. All the additional
consequences that result from the game have the potential to generate more interest from the
fans who will have an additional incentive to look at the game even if their favourite team is

\textsuperscript{13} For a complete discussion on the monopoly model see PINDYCK, R.S. AND RUBINFELD, D.L., 2005.
not implicated. This will result in a higher demand for the game coming from these fans interested in the impact of that particular game on the championship.14

Sports organisations recognized the potential of the market for championship early in the history of professional sport. It took different names such as Federation cups in soccer, pennants in baseball and the oldest one in North America, the Stanley cup in ice hockey. The idea was simple: teams would play a determined number of games in order to try to find out the champion. The original declaration made in 1892 by Frederick Stanley, The Lord Stanley of Preston, Governor General of Canada who created the Stanley Cup, showed that he understood the financial and sporting advantages linked to the determination of a champion

“I have for some time been thinking that it would be a good thing if there were a challenge cup which should be held from year to year by the champion hockey team in the Dominion [of Canada].

There does not appear to be any such outward sign of a championship at present, and considering the general interest which matches now elicit, and the importance of having the game played fairly and under rules generally recognized, I am willing to give a cup which shall be held from year to year by the winning team.”15

After the challenge or tournament cup, the league seems to be the next logical step since it is an extension of the cup concept. A league could be interpreted as a yearlong tournament in which teams would compete. Both forms of competition are only a structure used in order to exploit the market for championship. If the production of individual games seems to follow rules of its own, the market for championship is subject to the normal economic rules when sport organisations try to maximize their revenues. When exploiting the championship market, most professional sports move in the same direction in their search for profit: the monopoly. There is no need for a competitor to exploit a championship, like in any other business. Two leagues competing will have no incentive to collaborate since they fight for the same market shares. In none of the existing professional sports there is a competition between two leagues for the same championship market in the same territory. There were, in the history of most professional sports, competing leagues, but this always ended with the death of one protagonist or the merger of the competing leagues.16

---

14 The demand for professional sport will be examined in details in Chapter 2.
Although a championship can exist using a non-league context as the various challenge cups show us, the league facilitates the creation and management of the championship. With the league, the number and location of all participants are known at the start of each season. This allows for the creation of a set schedule for each team which results in some certainty in the planning of the activities of each team. With a schedule teams are, to a certain degree, able to start operating a business since their revenues and expenses become predictable. The teams know the number of games that they will play and the opponents that they will face, how far they have to travel, how many days they will be on the road, etc. The existence of the league allows the clubs to exploit in an efficient way this new good, the championship. As time passed, some organisations found new ways to do so by giving more teams a chance to win it by creating an elimination playoff tournament or meta-championships like the UEFA Champions’ league. Let us now have a more detailed look at the league, the most common way of exploiting the market for championship.

1.2. What do leagues actually do?

The definition of “league” is relatively simple and accepted, with some minor variations in wording, by most people studying that field

“(A) group of teams which play the same sport or activity in competition with each other.”\(^{17}\)

But that definition is so vague that we need a closer analysis to determine the nature of leagues. In that first reflection on what leagues do, I will try to go beyond the individual examples and try to look at the general concept of league: the common denominator of the idea behind the exploitation of the market for championship.

1.2.1. The “team goods” of sport

When they are created, leagues are given responsibilities by the founding teams. This coordination role is found, in one form or another, in all leagues around the world. These responsibilities could be qualified as “team goods” since they are common to all teams. In

addition, they are willingly delegated by the teams to the central authority of the league. Teams seem to have concluded that it is in their interest to centralise the management of these team goods. The first team good the league takes care of is the problem of finding opponents to play against and determine when to play against them. This was solved by the creation of fixed schedules in which all teams will face each other in turn for a certain amount of time. This predictability allows teams to plan the travelling expense, to make promotion and to develop rivalry between them. This in turn allows the teams to invest in a stadium where they will play half their games and attract a more constant crowd with a little development and marketing since the same location will be used for these games. The schedule also fixes a term for the championship to happen, creating an active and non-active part of the season. This off-season is crucial for teams as it allows them to have a quiet period to make modifications to their business and/or sport model.

The league also allows the teams to agree on the rules of the game that they are going to play. A minor variation in the rules may affect greatly the strategy of the sport and someone not used to playing with them may find it hard to compete. By making sure that everyone plays according to the same rules, the league prevents last minute rule changes which might benefit one team or another.

Competitive balance, a subject that will be explored in the next chapter, is another team good\(^1\). It means that the difficulty of the league’s task lays on how to reconcile the concept with the desire of individual teams to win games. Teams know that a winning record attracts more fans in their stadium and result in higher revenues\(^2\). In that perspective, teams do not want to adopt measures that would help other teams since a win for the competition is a loss for them. They will not want the others to be better than they are. The league must be the buffer that keeps the competitive balance in check, giving all teams a chance to win the championship and preventing the domination of only a few teams. Leagues usually do so by applying mechanisms like amateur draft, salary caps, revenue sharing etc. But the impact of these methods on competitive balance is uncertain and open to debate\(^3\). Chapter 2 will examine the question of competitive balance in more detail.

Finally, marketing is another aspect where incentives exist for the various teams to band together in a league to be more effective promoters of the championship. Promoting a team will result mainly in the exploitation of the market for games. Teams have the incentives to promote themselves in order to get more revenues. Only with a league it is possible to effectively promote the championship since this organisation is not bound by individual games but is there to promote the idea of the championship for the benefit of all participants. This is the kind of approach that will allow tapping on the demand for matches on television, stimulating the global demand for the sport and creating a brand name for the league and teams. Promoting the championship does not make sense for a single team since it allows its opponents to free ride on its investments. In that context, there are no incentives for an individual team to promote the championship and to concentrate their efforts on the promotion of individual games. In order to have the efficient amount of investment in the promotion of the sport, it is best to leave it to the league. The same logic seems to apply to the national television contracts; bloc negotiation seems to be able to yield more revenues for all the teams than when it has to be done individually since the league can use its monopoly power to attract more money from the networks. The ability to sell all the games in a package allows the league to sell the best and worst matches at their best prices which helps all teams maximize their revenues\(^{21}\).

Understanding the economic role of a league allows us to find even more common points between the different types of league and to look past the details and differences that may have arisen for purely historical reasons. It is therefore crucial to have a good economic understanding of what a league is. This is the purpose of the following section.

**1.3. The economic nature of sport leagues**

This section will look more precisely at the economic nature of leagues. With a good understanding of what a league is from an economic point of view, we are able to pin-point its role in the professional sport business model.

---

The actual consensus is to qualify sport leagues as cartels. However, there are other explanations that have been offered over the years on the nature of leagues and some of them are still worthy of study since the cartel model fails to account for certain aspects of the sport business. Determining the economic nature of sport leagues is of great interest for jurists since it allows them to see what kind of “creature” is regulated. It also allows a better perception of the unique features of the industry and if they require special legal attention.

1.3.1. Sport as a multi-plant firm in a Natural Monopoly market

The first proposition about the nature of leagues was made by Neale. He proposed that since leagues were the entities that have the real decision power in the production of sport, they were the “firm” of professional sports. The teams were only plants that produced the sporting good. This definition explains the importance of cooperation between teams since they are not, in that model, competitors but different plants of the same firm. Neale considered that the various characteristics of the sport market made the multi-plant firm the only viable form of exploitation of that market since the market for championship was, in his opinion, a natural monopoly.

He explained that the market for championship is a natural monopoly because it exhibits most of the characteristics of this market with high barriers to entry coming from the high costs associated with starting a league (stadium, network effect), high fixed costs (stadiums and players salaries) and low marginal costs (television broadcasting). In that context, the league represents the single monopolistic firm that is better placed to exploit the market efficiently. This theory finds strength in the fact that in history nowhere two competing leagues have survived for a long period in the same sport and territory. By the term “competing league”, one must understand leagues that are at the same level of talent and competing for the market of the top championship in the sport, in short, competing for the same market for championship. Leagues of various levels can survive on the same territory, as they do not

---

23 For more on the cartel theory see section 1.3.3.
26 ROSS, S., 1991, see note 16.
exploit the same championship since the lower level leagues openly recognise that their level of play is inferior to the top league. In the case of competing top leagues, the situation usually ends up with the demise of one league or the fusion of both of them\textsuperscript{27}.

But Neale’s concept of the multi-plant firm is not without its weaknesses. The main fault of this theory is that it assumes that the league is all-powerful and has complete control over its teams. In a multi-plant firm, plants are not given a lot of discretion; they generally follow the instructions of the firm. In most professional sports, it is the opposite: the league is the entity that is subordinated and the teams are the ones that give the instructions. The leader of the league is usually an employee of the teams since he has been elected by them. In addition, some of the most important decisions in the management of the league, like the entry of a new team or the signing of a new television contract, are subject to the approval of the teams\textsuperscript{28}. The teams seem to be the centre of decision even if the league officials have wide discretionary powers on some aspects of the sport business. So in theory the multi-plant model is attractive, but in reality the structures of the various leagues make it an inappropriate model for the industry. Still, the natural monopoly idea remains unaffected and is even today relevant to explain the nature of the market of professional sports\textsuperscript{29}.

1.3.2. The league as a joint venture

Neale’s paper started the debate on the economic nature of professional sport leagues and it opened the door to other theories. One of the most interesting theories that emerged, because of its legal implication, consists of defining the league as a joint venture between the various firms, the teams\textsuperscript{30}. In that model, teams are seen as independent businesses that are free to pursue their own economic advantage, but, like firms in a joint venture, agree to restrain their freedom of action in order to exploit the market for championship by joining or creating the league. It explains the fact that teams, while remaining fierce competitors on the field,

\textsuperscript{27} Idem.
collaborate off it and even adopt measures that could prevent individual teams from establishing a dominant position on the field in an effort to promote competitive balance. In that joint venture perspective the incentives to join the league come from the following trade off: the teams hope that the revenues coming from the added value of the championship on each game are greater than the loss of revenue in each game coming from the fact that teams relatively equal in their playing strength compete and that it may end up with a bad record. If we assume that teams rely on this trade off, it would explain their behaviour in joining leagues that can sometimes take decisions that are contrary to their individual interest.

Additionally, leagues have incentives to argue that they are joint ventures since this legal entity has the potential, in certain jurisdictions, of sheltering them from the application of certain laws. Most antitrust or competition laws contain a provision that exempts joint ventures or collaborative efforts between competitors that are related to joint ventures from the application of these laws. By allowing joint ventures, the law acknowledges that the consumer may sometimes benefit from the collaboration of firms or that some ventures would be unattractive if firms could not collaborate to benefit, for example, from economies of scale. One of these exemptions is the single entity doctrine in the US antitrust law.

1.3.2.1. The single entity doctrine

The leagues have always been reluctant to accept the intervention of the courts in their affairs and vigorously defended their autonomy. In the USA, when faced with antitrust litigation, leagues tried to use the “single entity doctrine” exception to the US Sherman act (US Antitrust law)\(^{31}\) that closely resembles a joint venture exception. This doctrine states that firms can be exempted from the application of the Act if they are able to prove that their operations are so intertwined that they can be seen as operating like one single entity.

This doctrine requires that the court, when examining the venture, looks if the concerned entity is fully-integrated and if the challenged decision (pricing, investment, contract, etc.) was taken by a single entity and not an agreement between independent economic actors. The fact that the individual firms compete in other markets and/or formerly competed in the products that are the result of the joint venture is irrelevant to the analysis\(^{32}\). The court will


\(^{32}\) *Texaco, Inc. v. Dagher*, 2006 WL 461525 (Feb. 28, 2006).
also look at the structure of the venture in order to see if it can identify “independent centres of decision-making”\textsuperscript{33} within it. If the court finds that the governance structure features more than one distinct centre of decision-making, it will reject the defence. The doctrine provides a second stage “actual or potential competitors” test\textsuperscript{34} that asks whether or not the parties to the joint venture contribute complementary assets, complementary capabilities or other complementary inputs. A negative answer to that question will deny the entity the right to use the single entity defence. This second stage is a \textit{de facto} rule-of-reason analysis that asks whether or not the restraints instituted within the venture are present across different products (horizontal restraints) or in the different types of processes linked to one product (vertical restraints). Finding that restraints are horizontal is the equivalent of finding that parties are \textit{not} contributing complementary inputs and that the parties are “actual or potential competitors.”\textsuperscript{35}

1.3.2.2. The single entity doctrine in sports

Sport leagues have tried numerous times to invoke the single entity defence. It started with the NFL in the \textit{North American Soccer v. NFL} case\textsuperscript{36}, but the court rejected the argument that the league constituted a single entity saying that the NFL is best defined as a collection of individually owned separate teams. The court added:

“\textit{[T]o tolerate such a loophole would permit league members to escape antitrust responsibility for any restraint entered into by them that would benefit their league or enhance their ability to compete even through the benefit would be outweighed by its anticompetitive effects.}”\textsuperscript{37}

But this did not seem to prevent the sport leagues from hoping to benefit from this exemption. On the contrary, they were comforted by the fact that the courts suggested that sport leagues could secure single entity status and acknowledged that teams could produce complementary inputs\textsuperscript{38}. The NFL tried no less than six times without success to plead that defence\textsuperscript{39} and no other league has been able so far to benefit from this defence, either. Courts have been

\textsuperscript{36} 670 F.2d 1249 (2d Cir. 1982).
\textsuperscript{37} Idem.
reluctant to grant it even in the case where a league structured itself specifically in order to try to benefit from this exemption.

1.3.2.3. The single entity doctrine in the EU

The joint venture economic model is of little help for European leagues since there is no equivalent to the single entity defence in EU law. The closest thing would be a Commission exemption using paragraph 81 (3) of EU treaty usually granted on a case by case basis. However, it is very unlikely that such an exemption could be granted to a league since the Commission usually sets as a condition for granting it the fact that the participants should not control more than 5 to 10% of the market in which they are operating. It is hard to see how a sport league would meet this condition since it effectively controls 100% of the market. One could try to encourage the Commission to define the market as the market for entertainment, but it is generally admitted that this definition is too broad and does not describe properly the market in which sport leagues operate. The only hope of sport leagues would be to lobby the Commission so that it bases “itself on considerations connected with the pursuit of the public interest in order to grant exemption under article 81 (3) of the Treaty.” But such an exemption would have nothing to do with joint ventures or the single entity defence.

1.3.2.4. The end of the joint venture

Ultimately, the joint venture model does not seem to be applicable or useful to sport leagues. It suffers from too many problems to be an efficient and robust model. Even in the jurisdictions where it is easy to apply, sport leagues have repetitively failed to convince the courts that they can benefit from this exception. In addition, this model became outdated because of the rise of a better explanation on the nature of sport league: the definition of sport leagues as cartels.

---

1.3.3. The cartel

It is generally agreed that the cartel is the structure that defines best what leagues are. The blueprints of the professional leagues were drawn at a period, late XIX\textsuperscript{th} century, in which antitrust laws did not yet exist. Businessmen of that time thought it was natural to set up a cartel with their competitors in order to control their respective market. Sport leagues constitute attractive cartels since they fulfill most of the conditions for a stable cartel: an organisation that allows the members to agree on prices and/or output and a market with an inelastic demand curve for both stadium tickets and the broadcast of games. The league itself provides a forum where the teams can speak to each other, negotiate and enforce agreements. Some agreements, like making sure that all the teams are playing by the same rules of the game, do not automatically have a competitive impact on the market. However, others like territorial exclusivity, the reserve clause, the amateur draft and limiting price competition by making the league the sole negotiator of the broadcasting rights of the league are clearly put in place in order to manipulate the market in favour of the members and to reduce potential competition between them.

Even if the cartel theory was formulated looking at the closed leagues of the North American market, it is also applicable to the European leagues since the only difference is that there is a small change in membership at the end of the year. Some European “open” leagues even offer a transition package to the teams that get relegated. In addition, rich teams are usually in control of the cartel and exploit it for their benefit since they do not have to fear relegation. One simply has to have a look at the way the money coming from the television broadcasting contracts is distributed with a skewed scale in favour of the most successful and not redistributed equally between teams.

---

46 PINDYCK, R.S. AND RUBINFELD, D.L., see note 13.
49 More on that subject later, see section 1.4.3.2.1.
1.3.3.1. The consequences of cartelisation of sports

As a cartel, the clubs can extract a monopolist rent in the form of high price tickets, high television contracts, high prices for merchandising items and subsidies for stadium construction or renovation from the governments\(^{51}\).

By limiting the number of teams and granting them local monopolies, the cartel protects its members from the competition they could make between themselves. A perfect competitive sport market would result in a situation in which big cities would have multiple teams until the market shares of all teams in the league would be relatively equal and all potential markets would be occupied\(^{52}\). This would lead to the consequence that the market would be saturated with teams and only the normal economic profit would be possible.

But a rational closed league will always leave a certain number of potential markets unoccupied. This allows the league to use the treat of transplanting existing franchises in these vacant markets to make the consumers accept their non-competitive actions and allow them to extract higher monopolistic rent\(^{53}\). Nevertheless, the local monopoly has some advantages since it allows the league to cross-subsidise some of the teams that would not survive in a pure competitive market. In such a market, the teams in smaller markets would either disappear or move to a bigger one. The local monopoly allows some fans to have access to the sport, access that would be denied to them in a pure competitive market. But the subsidisation of the smaller teams is not purely altruistic from the cartel point of view; it is good business, too. First, it allows for a relatively equal distribution of the clubs in the country, developing a national fan base for the championship. Secondly, it preserves the collective benefit of future expansions. When a team moves from a small market to a bigger one, the owner of the moving team is the sole beneficiary of that move: he gets higher revenues but does not have to share them with the rest of the league. If a new team is created in order to exploit the same bigger market, this new team must pay a fee to enter the league. This fee gets equally redistributed between all existing teams so that each of them gets part of the benefit of the exploitation of the new market. Finally, placing teams in all kinds of

---


\(^{52}\) ROSS, S., see note 16.

\(^{53}\) STEIN, B., see note 51.
markets prevents the formation of rival leagues. For a rival league, the production function may be different from the existing league and the small markets may end up being interesting and profitable. By making sure to have teams in all types of markets, the established league prevents the formation of various competitors in all these markets and secures the survival of its monopoly.

Cross-subsidisation is less present in the EU since the relegation and promotion system makes membership in the league unstable and reduces the efficiency of the cartel. As a result, teams in Europe have a tendency to concentrate in the bigger markets like London and Southern England for the Premier League, Northern Italy for the Serie A and Western Germany for the Bundesliga. Relegation and promotion results in a team territorial distribution pattern that is closer to the one in a perfectly competitive market.

The cartel is the best explanation available on how sport leagues operate. This explains the importance of antitrust law in the legal and corporate development of leagues, especially in the USA. Cartelisation can have some positive aspects for some consumers since it allows leagues to cross-subsidise small markets that would not survive in a purely competitive market. But this is clearly not sufficient to compensate for the negative aspects of the cartel, mainly the maximisation of the profits of teams, without any regard for the deadweight loss that they create in society at the expense of the fans.

1.3.3.2. Why not break up the cartel and let the market do the job?

As shown in the previous sections, leagues are a device to exploit the market for championship. Leagues are very directive in their approach; they determine the schedule, the venues and the procedure by which the championship is won. But is this interventionist approach the most efficient way to exploit that market or would the free market be able to do the same job or a better one?

We can have a partial answer to that question by looking at the situation that prevailed before the advent of the leagues in what was called the barnstorming days. In that type of market, teams used to arrange their own games; they sometimes travelled from town to town or stayed

---

in their hometown to face other travelling teams. In that type of organisation, one may argue, the championship is absent. Without a definite structure, it is hard to determine the best team in the sport since teams would meet different opponents of different strength and may even not meet each other. Multiple teams could have the same winning record. Such a situation happened in the early days of the Stanley cup in ice hockey. In 1894, four teams were tied for the best record and a playoff, the first in the history of hockey, needed to be organised to choose the challenger to the champion.\(^55\) So a structure is indeed needed to exploit the market for championships and the market seems unable to provide such a structure. Although a championship would be possible without a league, as the numerous soccer cups have shown, you need some sort of organisation to impose structures to exploit the championship. The triumph of the league in the market for championship is due to the fact that it is better at determining the champion than a single elimination tournament. In the later format, all hang on one game and some factors other than the quality of the team may distort the results. Over a longer period and multiple matches, like in a league, the dominant factor ends up being the quality of the team and not pure chance.

Another thing that the market seems incapable of doing is the standardisation of the rules of the game. There seems to be a need for a central authority which determines the rules and enforces them. There are numerous cases in history in which different types of games co-existed being played in different parts of a country. This made it difficult to determine the champion until some kind of unified rules were agreed upon. The most extreme case of such fragmentation of the sport happened in Australia where four types of “soccer” exist! The market seems to have encouraged diversity and today Australian Football, Rugby League, Rugby Union and Soccer are all represented by various professional and amateur leagues. One may imagine that if a league had, in the early days, merged all the codes into one, the resulting sport would have exploited the market for championship in that country more efficiently. Instead, we have four markets today for four different championships, a situation that makes it difficult for some leagues to survive\(^56\).

\(^{55}\)NHL Website: http://www2.nhl.com/hockeyu/history/cup/incidents.html, date of download March 2005.

\(^{56}\)Interviews with Mr. Alan Vessey, Football Services Manager of Football Federation Australia and Mr. Lawrie Woodman, National Coaching Development Manager & Game Development of the Australian Football League, March 2006.
So for structural and economical reasons, the league seems to be an efficient way of structuring the professional sports industry to exploit the market for championship. This market seems to be in need of some kind of structure to be exploited properly.

In the next section I will examine the various ways in which the leagues organised themselves in order to show that their differences are mainly a result of history and not of a fundamental difference between the two business models relegation system and closed leagues system. In conclusion, even if the differences seem important, fundamentally, both models do not differ much.

1.4. The genesis of professional sport leagues

1.4.1. Barnstorming

Different professional sports seem to have started more or less the same way, using what is referred to as the barnstorming team model. The professional teams travelled around the country and challenged the local teams to face them in a game. They would get part of the revenues that would result from the presentation of the game, move to another town and start the pattern again. As mentioned earlier in this chapter, if this model was enough to exploit the market for games, it was however not appropriate to exploit the market for championship, the most lucrative market in professional sports. Other types of organisations were therefore founded in order to exploit that market and the barnstorming teams slowly died out.

1.4.2. The tournaments

The next step in the organisation of professional sport was to create an independent competition in order to allow teams to compete for something more than just the result of one game. Instinctively, the early sport businessmen guessed that the market for the championship presented a lot of opportunities. In most cases, it took the form of a single elimination tournament. One of the first such tournaments was created in England by the Football Association (FA). The FA Cup, founded in 1863, is probably the oldest surviving tournament style competition. It is still held every year and can be won, in theory, by any English soccer team. Since its creation, it has been copied and implemented in many other countries.
1.4.3. The Leagues

The creation of leagues, with a fixed schedule that culminates with the team with the best record being awarded a prize, can be traced back to the creation of the baseball National League (NL) in 1876 by William Hulbert\textsuperscript{57}. He created the league together with a small group of businessmen who owned barnstorming teams. They agreed to play only against each other. In order to make the venture more profitable, the league granted territorial monopolies to the participants and provided a forum in which to collude with one another. In summary, it was monopolistic behaviour at its finest. In Europe, the NL model was imported in soccer, but European leagues tried to cooperate with the existing national associations that regulated the entirety of the sport, from the amateur level to the professional league, contrary to their American colleagues who created the NL to challenge the national association and proceeded to destroy any entity that could offer viable competition to their monopoly\textsuperscript{58}.

1.4.3.1. How professional leagues started

With the rising stakes coming with the creation of the market for championship, it was a logical step for teams to start paying players in order to be able to field the best team possible to win the championship. This was a change in culture since sport organisations were at their beginning mainly amateur entities that prided themselves on not allowing professionalism.

This was the result of the financial stakes in the market for games being too low and not providing enough incentives for teams to take measures to attract the best players. These incentives changed with the exploitation of the championship and the creation of leagues. It was only a matter of time before a gentlemen club would offer money to a good working class player who could not afford to play for free to quit his job and play as a professional. For such players, earning a living with their sports talent was a great alternative to the manual job they were likely to have. In 1871, the first purely professional baseball league was started in the United States. Soccer also started to recognise professionalism around that time and in 1888, the first professional league was founded in England\textsuperscript{59}.


\textsuperscript{59} SZYMANSKI, S. AND ZIMBALIST, A., see note 57.
But it was not until the beginning of the XXth century that leagues really achieved stability. Before that, it was not unusual for a league to end up its season with a lot less clubs than it had started with.

In the early leagues, predictability must not have been an important factor since it was hard to know if your opponents were going to financially survive the season. So the main incentive for teams to join early leagues and not simply keep on barnstorming must have been the exploitation of the championship market and the revenues perspective that it gave them.

In order to see the different methods in which different sports started to exploit and are today exploiting the market for championship, I will examine the model used by soccer in England and by baseball in North America.

1.4.3.2. England and the case of soccer

The organisation of soccer started in 1863 with the creation of the Football Association (FA) in London. It was created by a group of eleven soccer clubs with the initial purpose of establishing a common set of rules. Soccer, or at least a game implicating a ball, two teams and two goals, had been a traditional game in England since medieval times. It started to become more standardised in the early XIXth century when it was played as a winter complement to cricket in the public schools of England. Unfortunately, each school played by their own set of rules. The FA was created by graduates of various schools who wanted to continue playing their favourite sport after graduating. In order to do so they needed a unified set of rules. After much discussion, it was agreed that the FA would adopt the 1846 Cambridge rules, rules that were very close to the soccer that we know today.

The first real attempt made by the FA in the quest to take on the market for championship was the first “international” match- England v. Scotland in 1870 - and the creation of the FA Cup in 1871; a single elimination tournament open to all the member clubs of the FA. The original proposal for the cup was formulated as follows by FA Honorary Secretary Charles Alcock

---

Idem.

Idem.

FOOTBALL ASSOCIATION, official website. Available at: http://www.thefa.com/TheFA/TheOrganisation/Postings/2004/03/HISTORY_OF_THE_FA.htm, download date November 2005
"(T)hat it is desirable that a Challenge Cup should be established in connection with the Association, for which all clubs belonging to the Association should be invited to compete.\textsuperscript{63}

The Cup proved to be an immediate success both with the teams, with over 700 registrations in its first edition, and the public, with a record crowd of 2000 attending the first final\textsuperscript{64}. With the commercial success of the Cup, this gave incentives to the participating teams to take steps to increase their chances of success in the competition. One of the consequences was the first professional soccer players. Many complaints resulted from the fact that players were paid to play since this was prohibited by the FA rules. But it seems that this practice was tolerated and no serious sanctions were ever taken against teams that paid players\textsuperscript{65} until the FA finally embraced professionalism in 1885 by legalizing the use of professional players\textsuperscript{66}.

The popularity of the FA Cup prompted various regional associations in England to start their own cup. There was a multiplication of cups and the teams were more than happy to participate. However, the lack of predictability in the cup format caused problems to teams employing professional players. A team that got eliminated earlier than planned could face a financial crisis since it would not get the needed revenues to cover their expenses. In the idle time between two cups, a team could fold when having no revenues to pay its players until the next competition.

The Football League (FL) was created in 1888 in part to offer some kind of financial stability to professional teams. The inspiration for the FL was the baseball National League\textsuperscript{67}. The league did not destroy the cup format entirely; the FA Cup is still played today and is very popular with fans. But the league diminished the importance of the cup format for the teams. Revenues from a match against an obscure third league opponent cannot match the ones coming from a regular league game. In its second year of operation, the FL established two divisions and started the promotion and relegation system. The procedure was a little different than what it is today, but the basic idea had been introduced. This was another innovation

\textsuperscript{63} FOOTBALL ASSOCIATION, official website. Available at: http://www.thefa.com/TheFACup/TheFACup/History/Postings/2006/01/FACup_History.htm, download date June 2007
\textsuperscript{64} Idem.
\textsuperscript{65} SZYMANSKI, S. AND ZIMBALIST, A., see note 57.
\textsuperscript{66} Idem.
\textsuperscript{67} Idem.
copied from baseball’s National League\textsuperscript{68}. There are no clear explanations why this system did not survive in the USA but in England it seems to have been adopted as a compromise between the FL and the FA, the first wanting a commercially viable league, the second wanting to promote the development of soccer and the founding of new teams\textsuperscript{69}. After that, the model stayed pretty stable apart from some adjustment, like the 1922 expansion to four divisions, until the establishment of the Premier League in 1992.

This newer pyramidal relegation and promotion model is now the dominant one in professional sports and has been adopted in the majority of the sport leagues outside of North America.

For almost a century, the structures of English soccer remained the same but television and the revenues it could bring forth disturbed the existing structures. Before the advent of the Premier League, television revenue was shared almost equally between all the clubs of all the four divisions. Not surprisingly, the bigger, more successful first division teams did not like to share the spoils with the teams in the lower divisions that had no television exposures. Thus in 1985, the five biggest English teams formed an alliance to force the FA to give them bigger shares of the television revenues. They threatened to form a breakaway league if their demands were not met. Furthermore, English soccer was in complete disarray in the late 80s with rampant hooliganism and the Heysel and Hillsborough disasters that resulted in sanctions against English teams and their exclusion from European competitions\textsuperscript{70}. The pressure on the FA was too great and an eventual break away league could have destroyed its remaining influence and credibility. To prevent that, it accepted the demands of the big clubs and allowed the creation of the Premier League\textsuperscript{71}. This new league permitted the member clubs to keep all the revenues of the broadcasting contracts to themselves, a situation that they justified by the need to overhaul their stadiums to prevent further disasters\textsuperscript{72}. Additionally, teams set up protections against relegation and fire sales by creating a two years “parachute payment” that would help finance the teams that were relegated in the lower division and help

\textsuperscript{69} SZYMANSKI, S. AND ZIMBALIST, A., see note 57.
\textsuperscript{71} CONN, D., see note 70.
\textsuperscript{72} Idem.
them come back into the Premier League. This created in fact a semi-closed league, a hybrid between the traditional European league and the US ones. This is the form the Premier League still has today and the economic environment in which professional soccer players are evaluated. Although the semi-closed model is unique to England, most of the top European leagues are somehow similarly structured.

1.4.3.3. Major League Baseball

The first embryo of organisation in baseball can be traced back to 1858 when the clubs of the New York City area banded together to create the National Association of Base Ball Players (NABBP). Created by baseball aristocrats, the NABBP had little in common with a league and seemed more concerned with maintaining the class system inside the sport than accommodating any type of commercialisation of it. The goal of the NABBP was to regulate the behaviour of players: the first regulations they enacted concerned the banning of drinking and fighting during games, prohibiting betting on games by players and banning the compensation of players. Baseball teams were originally created by well-off gentlemen wanting to play the game with its regulations. The NABBP aimed to preserve their way of playing the game making sure that everyone knew his place and did not try to break away from the standards of society. But the exponential growth of the game in the following decade, over 300 clubs all over the US were NABBP members in 1867, made it impossible and unrealistic to exert the control that the gentlemen in New York would have liked on the game. This is when the commercialisation of baseball started, mainly with teams charging admission fees for games and the remuneration of players. This lead to a conflict with the NABBP and in 1871, the National Association of Professional Base Ball Players (NAPBBP) was created for teams that wanted to run baseball like a business. Although it survived only five years, the NAPBBP had a lasting effect on the world of baseball since it invented the notion of national championship in the USA. Although there were no fixed schedules and no league structure per se, the NAPBBP was the first organisation to try to exploit the market for baseball championship by awarding a prize at the end of the year to the NAPBBP team with the best record. But the lack of structure, fixed schedule and stability in membership doomed...
the NAPBBP that folded in 1876\textsuperscript{76} to be replaced by the first real professional league in the modern sense of the word, the National League.

1.4.3.3.1. The birth of the National League

The creation of the first baseball league was more an accident than anything else: it was not brought about from a visionary and noble purpose but was simply created to accommodate the needs of one team wanting to break away from the NAPBBP. At the start of the 1876 season, William Hulbert, president at the time of the Chicago White Stockings, had a problem: five of his star players were threatened to be banned from the NAPBBP because Hulbert had signed them to his team using questionable methods. Faced with the possibility of sporting and financial failure, Hubert decided to start his own organisation to exploit the championship market, the National League (NL). The NL was a totally new way of organising and structuring a championship since it was a closed league that would play all its game between the member teams. Hulbert’s inspiration was the cartels that were very common at that time in different industries in the US. Each team was awarded a territorial monopoly, the exclusive rights to exploit the NL market in a given city. This was an important innovation that still differentiates leagues on both sides of the Atlantic today. As for the team structures, they were inspired by the corporations, with presidents, chairmen and boards of directors. The owners perceived themselves as totally controlling the means of production, including the players. This was in conflict with the vision of sports at that time since a significant number of clubs in the NAPBBP were cooperatives of players that shared revenues. Additionally, the players had representatives in the various organs of the NAPBBP. So the NL took the players out of the equation. With this new league, they were simply chattel and only the owners were allowed to have a say in the decisions of the NL.

What is today known as Major League Baseball is the result, more or less, of this vision and is totally independent from any national or international baseball association. There is an International Baseball Federation (IBAF) and a national federation (USA Baseball), but they do not have any jurisdiction over the professional players playing in the MLB\textsuperscript{77}. Also, there are no international transfer cards or accreditation. The only purpose of the IBAF is to

\textsuperscript{76} Idem.

\textsuperscript{77} USA Baseball was created in 1978, more than 100 years after the birth of the National League. Source: USA BASEBALL. Official Website, available at: http://mlb.mlb.com/usa_baseball/about.jsp date of download October 2005.
organise tournaments outside of MLB implicating national teams mainly composed of amateurs or marginal professional players. It does not have any power to discipline MLB’s clubs or players. MLB basically does as it pleases, enjoying the power vacuum created by the absence of a national federation in the US and the insularity of the sport, being played mainly in the United States and therefore not needing any form of international coordination.

1.4.3.3.2. Closed leagues

The basic model adopted by the NL was copied by all North American Leagues and developed into what is now called the closed league model. The main feature of the North American leagues is that they are closed to outside participation: once the league is created, there is no possibility for any outside team to access it unless the existing members allow it via what is called an expansion. In history, teams tried to force their way into the system via lawsuits, but the courts refused to force leagues to allow new members in\(^78\). So unless there is an expansion or a financial failure from one of the existing teams, the composition of the league will stay the same. Once you are in, you stay in but if you are out, you have to beg your way in.

Leagues in North America are independent entities. There are no national or international federations that have a say in the way the league runs its business; only the league can take decisions that would affect the status of a player or a team. North American leagues are in a world of their own and independently structure the ways in which they exploit the market for championship. The early influence of corporatism is reflected in the decision structures of the leagues. Each team is considered a shareholder in the league and has a say in the decision making process. This system usually follows the “one team, one vote pattern”, each team is considered to be the owner of one share of the corporation that is the league and this share gives it the right to vote on decisions taken by the league. These decisions usually range from changing the rules of the game, expansion, allowing ownership change, introducing new technology, sponsorship contracts to electing the commissioner, enacting the collective bargaining agreement, etc. Like in ordinary corporations, decisions are usually taken using a simple majority but the statutes of the league may ask, in certain cases, for a super-majority, especially if the decision has a lasting impact on the league. This way of proceeding serves some leagues well but in some cases, decisions tend to be taken using the lowest common

\(^{78}\) *Mid-South Grizzlies v. NFL*, 720 F.2d 772 (3d Cir. 1983).
denominator or looking for short term gains without a long term vision. Owners tend to think about their individual gains rather than trying to maximize league revenues and without strong leadership or vision, solving this selfish vicious circle becomes difficult. The task is even more complicated since the head of the league, usually called the commissioner, is in fact an employee of the teams that are members of the league. They usually elect him like the shareholders do with a CEO. So the capacity of the commissioner in imposing his views to the owners can vary greatly depending on his persuasion abilities, the powers delegated to him by the statutes and the willingness of the owners to be lead. Although the powers delegated to him can be wide and sweeping, commissioners have the tendency to interpret them very restrictively not to end up on the wrong side of an argument with team owners who could easily make them lose their jobs. Some commissioners are more or less mouthpieces for the teams like Allan Huber "Bud" Selig Jr, Commissioner of the MLB, whose family still owns an MLB team. Other commissioners, through sheer personal strength of character, ego, powers delegated by the statutes or alliances with club owners, become the defenders of the interest of the whole league against the selfish influential owners. The ultimate goal of the commissioner of a league should be to make the teams accept solutions that may not be revenue maximising for an individual team but that will maximize the revenues of the league.

1.5. The coop league?

Even if North American leagues were inspired by the corporate model, they may have in fact more in common with a cooperative. All league wide decisions must be voted on by the different teams and they are the central source of power in the decision making process. Additionally, the votes are not linked to shares but to teams. However, in the management of the internal affairs of their teams, they remain mostly autonomous entrepreneurs; one could say that the North American leagues are partial vertical forward-integrated entities used by all team owners to exploit the market for championship. In layman’s terms: the league is a vehicle in which the individual teams delegate part of their powers and production capacities.

79 SZYMANSKI, S. AND ZIMBALIST, A., see note 57.
80 LENTZE, G., see note 28.
82 Idem.
European leagues are more loose contracts between various teams and the national federations. There are no centralized power structures for the league and no commissioner figure. It is difficult to see any integration apart from the minimum necessary to make the league work. Such arrangements are more in the line of a pure cartel.\(^{83}\)

As discussed earlier in this chapter, leagues are usually qualified as cartels, but as we have seen the situation may be a little more subtle and some aspects of the business may well not be according to a cartel. They have some characteristics of coops as seen previously but also share some with corporations and joint ventures. It therefore makes more sense to qualify a league as a hybrid organisation\(^ {84}\). That is why it is reasonable to judge every action individually and not automatically label them as harmful and the act of the cartel. Such an approach allows us to see that some actions may not be anticompetitive and therefore may have some benefit or at least have no harmful effect.

**1.6. Conclusion**

We have seen that the rise of the different sport organisations was a result of a desire to exploit the market for championship, the market in which there is the most revenues for them and all the other parties involved in the industry.

We also have seen that even if there seems to be numerous good economic explanations that have been proposed for leagues, the only one that seems satisfactory is to see the league as a cartel. But we have seen that it is a special cartel, one in which some decisions are beneficiary to society in general and would not happen in a competitive market.

What does that mean in the context of this thesis? It shows that accepting the “special” status of sport and letting leagues regulate themselves like they want is not automatically harmful for society. But one must remain careful since, for other actions that are not specific to sport, the cartel is more likely to be harmful to society. The development of the leagues over time has created complex entities that are difficult to categorise and must be approached carefully. Although courts have been reluctant to do a separation of the different characteristics of the sport industry, this does not mean that it may not be efficient to take that approach.

\(^{83}\) Idem.  
\(^{84}\) Idem.
However, for the purpose of this thesis we need to understand how leagues are structured and how that can affect the player. What we have seen in the past chapter is that, independent of its structure, the league is an instrument to exploit the market for championship. By exploiting this market, the sport industry has access to more revenues and the players therefore have to position themselves in order to benefit from it.

So this debate on the nature of sport organisations allows us to see that the players are mainly evolving in an economic environment where their employers are firms operating a cartel. However, some actions of the cartel may be to their benefit because of the nature of the industry in which they work. One must therefore remain careful and examine, for the purpose of this thesis, each action individually.

In the following chapter I will examine one of the special aspects of the industry, one in which the actions of the cartel have the potential of giving positive results to society: competitive balance.
Chapter 2 – Competitive Balance or Uncertainty of result

What’s the most you ever saw lost on a coin toss?
Sir?
I said what’s the most you ever saw lost on a coin toss.
Coin toss?
Coin toss.
I dont know. Folks dont generally bet on a coin toss. It’s usually more like just to settle somethin.
What’s the biggest thing you ever saw settled?
I dont know.
Chigurh took a twenty-five cent piece from his pocket and flipped it spinning into the bluish glare of the fluorescent lights overhead. He caught it and slapped it onto the back of his forearm just above the bloody wrappings. Call it, he said.

- Cormac McCarthy, No Country for old men

2.1. The need for uncertainty

The importance of uncertainty seems to be something that was instinctively understood by early teams and professional sportsmen, but it was not until the 1960s that its importance was actually theoretically demonstrated for professional sports.85 The goal of this chapter is to give the reader a clear understanding of this important concept. It drives the sport business, therefore it is crucial to master the concept and see its impact on the market for players. Before starting I would like to point out that in the sport business field, the terms uncertainty of results and competitive balance are used without distinction to define the same thing: the fact that every team should have a realistic chance of winning a sporting contest, be it a game or a championship.

2.2. The Louis-Schmeling paradox

Neale was the first to offer an explanation for the importance of uncertainty of result 86. To illustrate his point, he used a legendary heavy weight boxing match up.

Joe Louis was an American heavy weight boxer and Max Schmeling was a German boxer in the same weight category. They had fought the first time in 1936 in a match that was won by

---

85 NEALE, W.C., see note 24.
86 Idem.
Schmeling. In 1937, Lewis became champion of the world. When it became time to choose a challenger for Lewis’ next fight, it was generally agreed that Schmeling was the most “interesting” contender. He had triumphed over the champion once which had been Lewis’ only defeat so far\textsuperscript{87}. Schmeling was therefore considered the only one who could defeat the champion and every boxing fan wanted to see if he would be able to repeat the feat. To Lewis, Schmeling was the most dangerous contender, the one who had the most chances of defeating him, so in a sense it was a bad choice since he could lose his title to him. However, he was also the contender who sparked the most interest. A fight between Lewis and Schmeling would generate a high demand for stadium tickets and radio broadcasting; therefore it would be the match that would generate the highest revenues and the highest prize money for Lewis. In boxing, the prize money given to the fighters is correlated to the expected demand for the fight and this demand is greatly influenced by the excitement generated by the fight\textsuperscript{88}.

Neale correctly stated that excitement in boxing is related to the uncertainty of result of the fight. The less predictable the outcome, the more interest it generates and therefore the more revenues the promoter will get and in return offer a higher prize to the fighters in order to get them to agree to the fight.

The paradox in that context, seen from the point of view of the champion, is: the greater his chances of losing the fight, the greater his revenues are going to be.

In a more general sport industry sense, it may also be stated as such: a close sporting contest with evenly matched opponents will result in a greater demand for it on the part of sport fans that are willing to pay more to see it than a more predictable contest. Therefore, the contest will generate higher revenues for the contestants. In summary, uncertainty of result is the basis of the sport business\textsuperscript{89}.

In a perfect world, the two opponents should be of equal strength, having an even chance of winning the contest. This total uncertainty is what enflames the passion of the various fans and creates greater demand and revenues to all involved.

\textsuperscript{87} RING MAGAZINE, (1938). Louis’ only setback, July 1938. For the benefit of the reader, Lewis’ career record would end up being 68 wins and only 3 defeats. The Schmeling’s defeat was the only one he suffered during his prime. Lewis would stay undefeated from 1937 to 1949.

\textsuperscript{88} NEALE, W.C., see note 24.

\textsuperscript{89} Idem.
If the principle is pretty straightforward in an individual sport like boxing, in the case of sport leagues the equation and the application of that principle are more complex. As written by Judge Rehnquist in his dissenting opinion in *National Football League v. North American Soccer League*90

“Although individual NFL teams compete with one another on the playing field, they rarely compete in the market place.”

From these facts results a tension between the league and its member teams. On the one hand, the league wants uncertainty of result between the teams in order to create more demand for the sport and therefore more revenues. The individual teams on the other hand want to win every match and be the best to satisfy their fans who want to root for a winning side91.

In a league, the advantages for the individual teams from having equally matched opponents are not as clear as in the case of individual sporting contests. So it is up to the league to preserve that uncertainty of result for the good of all teams. As mentioned in chapter 1, competitive balance should be considered a “team good”.

Yet, the Louis-Schmeling paradox is not perfect and is unable to explain some aspects of the sport model. Later development in the field also revealed some other ways in which the demand for sporting events is influenced. This is what I will examine in the following section.

**2.2.1. Holes in the paradox**

The problem begins with the first underlying assumption: that the fans are homogeneous in their demand for what they want to see in a sporting contest. By saying that fans crave for uncertainty one assumes that all fans are alike and that they are all interested in seeing a close sporting contest. In reality, fans are rarely homogeneous in their demand for sport. A more realistic way of structuring the demand of fans would be to say that there are two types of fans: fans of a team and fans of the game92.

---

91 HOEHN, T. AND SZYMANSKI, S. see note 19.
92 Idem.
Fans of a team are those supporting one of the two teams involved in the contest. In consequence they are willing to pay more for a contest in which their team may have a better chance of winning. They typically derive high utility from a high-scoring game, won by their favourite team, which is not too one sided\textsuperscript{93}. A loss is not something that will be enjoyed and cannot be compensated by the fact that the opponents were evenly matched.

Fans of the game are spectators who do not root for any of the teams implicated in the contest; they only want to see a good close game\textsuperscript{94}. They are also the ones interested in seeing a nice display of skills from either a team or a player. For these fans, the winner or the loser has little effect on the utility they derive from the game. What does affect their utility is the quality of the game. It needs to be uncertain, closely contested and with a nice degree of spectacular displays of skills. These fans are the ones reacting in the way predicted in the “Louis-Schmeling” paradox. They have been gaining more importance with the advent of television and radio broadcasts: these mediums bring the game to fans that do not have teams in their own hometown or are unwilling to go to the stadium. Although less fanatical in their devotion, they constitute an important marketing section and are the reason why major networks are willing to pay a lot of money for the right to broadcast matches. This helps explain why uncertainty of result always seems to be in the mind of leagues searching for more revenues.

This heterogeneity, at least in these two types of fans, is represented by the demand for stadium admission at games and the demand for the broadcasting of games. The two types of demands react differently to the “Louis-Schmeling” paradox. According to studies of the demand in professional sports\textsuperscript{95}, the most important factors to explain stadium attendance are the chances of the home team winning the game and the importance of the game in the context of the championship. Therefore, the highest attendances were achieved in games in which the home team had good chances of winning and/or that had a great impact on determining the team’s promotion/relegation/European cups participation chances.

\textsuperscript{94} HOEHN, T. AND SZYMANSKI, S. see note 19.
So according to these studies, uncertainty of result has less of an impact on stadium attendance than the classical sport economics would lead us to believe. In fact, uncertainty of result is negatively correlated with the chance of the home team winning the game, meaning that there will be fewer fans present. So uncertainty of result has a negative impact on the first preference of stadium attending fans, the probability of a win by the home team. This also explains the resistance of individual teams to measures taken by leagues in order to promote uncertainty of results. Teams know that these measures are likely to negatively affect attendance in their stadium and result in lower direct revenues for them.

As to the second preference - the importance of the game in the championship - competitive balance is positively correlated with it. If the teams in the championship are evenly matched, this will create more games with an impact on the championship since the rankings are going to be very close. With each place in the championship closely contested, numerous games will have an impact on it. So the “Louis-Schmeling” paradox is not totally absent from the equation determining the demand for stadium attendance, only its effect is weaker than the original model would predict. Using the vocabulary of chapter 1, one could say that uncertainty of result affects stadium attendance in the market for championship only. In the case of the market for games, attendance is negatively affected by uncertainty of result.

Nevertheless, stadium revenues are now being eclipsed by another source of revenues on which uncertainty of result has an even stronger effect: the broadcasting of games on television. Revenues coming from the sales of broadcasting rights of matches are nowadays, directly or indirectly, one of the main sources of revenues growth for the teams and leagues. The National Football League (NFL) in the US is probably the best example. The latest contracts will give clubs around 60% of their revenues with an average of 3,05 milliards\textsuperscript{96} US dollars per year over six years\textsuperscript{97}. European soccer is not to be left out of that trend; the English Premier League has negotiated new contracts for the television broadcasting that rivals the NFL and are worth an estimated 3,5 milliard US dollars a year\textsuperscript{98}. Even if some other

\textsuperscript{96} One thousand millions or 1 000 000 000.
leagues have contracts that are not as rich\textsuperscript{99}, they still have an incentive to take good care of the television audience since it represents a source of revenues where growth is possible at very little costs. Attracting additional fans in front of the television screen causes almost no costs for the league and teams since the infrastructure needed to broadcast to 100 000 or 10 million persons is the same. In contrast, building additional seats in a stadium represents a considerable investment. In economic terms one would say that the marginal cost of accommodating additional television spectators is close to zero, so more demands in that market result in more profits for the league and teams and any measure that will enforce growth in that market is highly attractive. The demands of television spectators are therefore more likely to have an influence on the measures taken by the sport organisations to increase their revenues. As said earlier, television fans are more likely to be fans of the game and are positively affected by competitive balance in both - the markets for game and championship.

As already mentioned, sitting in front of the television requires less effort and investment than going to the stadium. The television watching fans, being less interested in the result of the game, are casual observers of sport who may have no preferences of clubs or are unable to see their favourite team on television due to commercial reasons. Since we can assume that these fans provide the majority of the people looking at the games on television, the broadcaster will pay more for a sport that will attract them and offer the exposures to advertisers. So once again, uncertainty of result takes the front seat of the economic development of sports if we assume that the fans of the game compose the majority of the television audience and should have a considerable impact on its size.

Television broadcasters determine the price to pay for the rights to broadcast the games based on the price they can charge advertisers to reach this public. The more people look at the game the more advertisers will be ready to pay for an advertisement during a sport contest. So in the case of television audience, the influence of the “Louis-Schmeling” paradox is stronger: the closer the contest, the greater the demand.

2.3. What is uncertainty of result in a professional sport league context?

The concept of uncertainty as presented in the one game context in which the “Louis-Schmeling” paradox was first formulated is relatively easy to understand. So far, I have discussed competitive balance and uncertainty without defining the concept. To paraphrase Bob Costa, an American sports writer “We may not be able to define it, but we sure recognise it when we see it”\(^{100}\). The goal of this section is to produce a reasonable definition of competitive balance in a league context.

2.3.1. Game competitive balance

The game competitive balance is mainly a measure for the difference in strength between two teams facing each other for one single match. The perfect situation, from a revenue maximising point of view assuming that greater balance brings greater revenues, is when both teams have an equal chance of winning the game. The closer it comes to a 50 percent chance of winning for both teams, the more balanced the game will be. This is the situation that was examined by Neale when he formulated the Lewis-Schmeling paradox. In the simple one shot exploitation of the market for game, competitive balance can be simply defined as: an equality of chances for both teams to win the game.

2.3.2. Championship competitive balance

This is when the problems start: in order to efficiently exploit the market for championship, the league also needs to give every team a chance of winning the championship. The league wants to maintain the suspense as to which team will be the champion/promoted/relegated/or will make the playoffs until the end of the season. This form of uncertainty is not focused on the results of individual games but on the addition of all the games of the schedules of all the teams. This multiplication of variables makes it difficult to have a simple definition of competitive balance like it was the case in the market for games. In the following section, I will examine three ways of determining competitive balance in a league.

2.3.2.1. Single season standard deviation ratio

One of the most common methods to find the competitive balance of a league is to compare it with the perfectly balanced league. In such a league, every team would have an equal number of victories and defeats. This reflects the fact that in each game, each team has an equal chance of winning. The reason why most studies have done such analysis is to find the standard deviation of the professional sport league. The closer the league is to the standard deviation of the perfect league, the better the competitive balance. This has been called the single season standard deviation\footnote{SCULLY, G. W., 1989, \textit{The Business of Major League Baseball}, Chicago: The University of Chicago Press and QUIRK, J. AND FORT, R., 1992, see note 22.}. It has the advantages of being simple to execute, relatively robust and deliver clear results. Nevertheless, it occults the individuality of the teams. The league may be balanced on paper, but this method will not tell us for example that the championship is constantly being won by the same team year after year. Additionally, if changes in the competitive balance are detected over the years, the analysis will not be able to explain why these changes occurred. The domination of one team in a league over a couple of seasons will surely affect negatively the measure of competitive balance; however, is the domination due to a disproportion of revenues or to innovations? Some big changes in the leagues like the signing of a new Collective Bargaining Agreement (CBA) may be wrongly interpreted as promoting competitive balance when in fact they are the trees that hide the forest and that change actually comes from a complex combination of lesser known factors. The single season standard deviation model tells us how the competitive balance is affected but not why.

Another final criticism is that this is a measure of competitive balance \textit{a posteriori}. It helps us see the level of equality in the past season but cannot help us predict the future.

I believe that the best way to understand the problems of this method is to take an example.
Let us look at the standard deviation ratio of the National Hockey League (NHL) for the years 2000 and 2003\textsuperscript{102}.

<table>
<thead>
<tr>
<th></th>
<th>2000</th>
<th>2003</th>
</tr>
</thead>
<tbody>
<tr>
<td>NHL</td>
<td>2.16</td>
<td>1.70</td>
</tr>
</tbody>
</table>

A perfectly balanced league would have a single season standard deviation ratio of 1. So the closest the single season standard deviation is to 1, the better the competitive balance of the league.

This data seems to confirm the allegations of the NHL that the rule changes implemented in the 2000 season helped the competitive balance of the league. Someone looking for an explanation could say that the introduction of two referees on the ice and various other rules changes implemented in these years had an impact on the competitive balance of the league\textsuperscript{103}.

But there may be another explanation for that competitive balance change. Between 1998 and 2000, the NHL added four new teams to the league (one in season 1998-1999, one in 1999-2000 and two in 2000-2001)\textsuperscript{104}. Expansion teams are usually weaker than established teams since they have to make their squads out of lower quality players than the average of the league. Therefore, these teams usually have a lot of catching up to do, especially regarding the creation of their development teams. This results in a team that is weaker than the average of the league\textsuperscript{105}. The presence of four such teams in the league in 2000 could explain why the single season standard deviation ratio to the perfect league is so high. These four weak teams had a big differential between games won and lost during their first years. In 2003, as all these teams had climbed up from mediocrity to average performance, they helped bring the ratio down and make the NHL look like a league with a better competitive balance. This seems to be a more logical explanation than rule changes. However, one cannot be certain since this method does not allow to see why the ratio changes. One could make a compelling argument


\textsuperscript{104} NHL WEBSITE: http://www.nhl.com/history/062597.html date of download: March 2006.

that the rule changes are at the origin of the change and it would be hard to completely
discredit it. To determine the cause, we need to use other tools and to look beyond the simple
ratio or try to use another way of measuring the balance of the league.

2.3.2.2. Championship variance balance

Another way to measure competitive balance is rather simple: one looks at the number of
times a team won the championship and if a league shows a great diversity of champions,
then you could say that it has a great competitive balance. This method has the advantage of
giving a relatively clear result that can be understood by anybody without the need to know
arithmetic or statistics. However, it does not take into account intra-season balance. This
means that this measure is incapable of showing if the champion was determined at the
beginning, middle or end of the season. A team that won three very closely contested
championships will give the impression that the league is not competitive and one where three
different teams won boringly at the middle of the season without much credible opposition
will show great competitive balance. This measure is helpful but too crude and imprecise to
accurately measure the competitive balance of the league. It still remains useful if used in
conjunction with other methods like the single season standard deviation ratio, but suffers
from the same problem as this method since it measures competitive balance *a posteriori*.

2.3.3. The Blue Ribbon Panel on Baseball Economics definition

As the previous sections demonstrated, it is difficult to find a satisfying statistical method that
provides a simple and satisfactory definition of competitive balance in a sports league in the
market for championship. Even if we can figure out an indicator for it, there are no guarantees
that it would be helpful for the development of the business model. Since we are approaching
the problem from a legal perspective, the desired definition does not need to be expressed
mathematically. What we need is a definition that will help understand the influence of
competitive balance on the league in general and on the market for players in particular.
Fortunately, there may be such a definition provided in the Blue Ribbon Panel on Baseball
Economics report\textsuperscript{106}.

download January 2005.
This panel of various experts was created by the MLB Commissioner in order to find ways to solve the crisis Baseball faced at the end of the 1990’s. The cancellation of the 1994 season due to a player’s strike triggered a fall in both stadium attendance\textsuperscript{107} and television ratings\textsuperscript{108}. Revenues of teams dropped accordingly\textsuperscript{109} and numerous teams faced financial difficulties. This forced MLB to reflect on the way it conducted its sporting and business activities.

Over the years, MLB left the impression that it had evolved into an unbalanced league in which only a limited number of rich teams got a chance to contest the championship\textsuperscript{110}. The single season standard deviation method also showed that the competitive balance of MLB had worsened between 1970 and 1990\textsuperscript{111}. In addition, studies demonstrated that there was at that time a strong correlation between the size of the market and the performance of teams in the championship. The championship was constantly won by the teams with the biggest salary masses\textsuperscript{112}. The Blue Ribbon Panel was then created in order to help MLB find a solution to its diminishing revenues. The league thought at that time that there was a link between the diminishing revenues and the imbalance of the championship. After numerous consultations and studies the commission wrote in its report numerous recommendations on how to fix the various problems of MLB. These recommendations and their potential impact on competitive balance would make an interesting study, but that is not the purpose of this thesis. What makes the report relevant is the way in which the panel defined competitive balance. It offers the following definition

“[P]roper competitive balance should be understood to exist when there are no clubs chronically weak because of [the league’s] structural features. Proper competitive balance will not exist until every well-run club has a regularly recurring reasonable hope of reaching postseason play.”\textsuperscript{113}

(Emphasis in original text)

As mentioned in chapter 1, in baseball the post season is a tournament that allows the eight best teams to compete for the championship. The competitive balance definition of the panel focuses on the ability of the teams to participate regularly in this tournament.

\textsuperscript{109} Idem.
\textsuperscript{110} COSTA, B., see note 100.
\textsuperscript{111} MERRIGAN, P. AND TRUDEL, P., see note 102.
\textsuperscript{113} LEVIN, R. AND ALS, p.5, see note 106.
The post-season is important in the revenue function of North-American professional teams since revenues earned during that period constitute pure profit. At this point of the season, the main cost aspect of the teams, the salaries of the players, has already been paid since the contracts cover only the regular season. The players generally get a bonus for their participation to the post-season, but that is less than what they receive in the regular season. For this reason, the post season can mean the difference between an annual profit and a loss for teams.114

For the Blue Ribbon panel, it seemed important to give every team the chance of taking part in that tournament. Although this could seem a definition that applies only to North American sports, it is very easy to transpose in the European context. In recent years, participation in the different European cups has taken more and more attention and importance in the various soccer leagues. In fact, these cups could be seen as playing the same role as the postseason for the North American teams, not because of the players playing these matches “for free” but because of the high revenues resulting from these matches.115 Participation in European competition is often seen as a reason for the imbalance between teams in the various national leagues since clubs use the European revenues to subsidise their national championship performance.116 For these teams, failure to reach European competition can have a disastrous impact on their finances comparable to the exclusion from the postseason for a North American team.

Therefore, I think that this is the most useful definition for the context of this thesis. It is result oriented, focusing on a very clear and well defined objective. In addition, it allows taking into consideration the now very important commercial aspect of sport and the fact that the additional revenues coming from the participation in this extra tournament are also vital for clubs in order to keep competitive balance in the league. Like a virtuous circle, the more diversity in the teams that are allowed to participate to these tournament, the more the revenues are going to be equal inside the league giving each team equal capacity to invest in players. So for the purpose of this thesis, competitive balance will be defined as follows:

114 Interview with Mr. Julien Brisebois, Vice-president to Hockey operations of the Montreal Canadien, March 2007
115 DOBSON, S. AND GODDARD, J., see note 47.
116 SZYMANSKI, S. AND ZIMBALIST, A., see note 57.
117 Idem.
Every well-run team in the league has a regularly recurring reasonable hope of reaching postseason play or European competition.

2.4. Competitive Balance and its effect on players

In chapter 5, I will look in detail at some of the measures taken by leagues to influence the market for players to allegedly promote competitive balance. In professional sports, players have mostly been seen as agents of imbalance and leagues and teams have usually taken measures to prevent that all the best players end up in the same team.

But for the moment, we must reflect on the impact of competitive balance on the value of players, assuming that there are no restrictions imposed on them. As stated earlier, competitive balance usually results in greater demand for the sport and that in turn will result in more revenues for the teams. With their employer earning more revenues, the players will be able to ask for higher salaries and therefore benefit from the competitive balance as well.

In addition, if the championship is closely contested and assuming that the teams are not purely profit maximising entities, it is very likely that the demand for the service of players coming from the clubs also increases. With each team trying to get an advantage on the playing field, there will be a competition to secure the services of the best players. Some players may benefit more than others but the aggregate value of the players will grow and all players will benefit from the higher competition for their services. As every team fights to get the edge on its competitors, every player assumed to help it in the championship is likely to be seen as more valuable.

In short, competitive balance is also likely to influence the value and revenues of players positively.

But it is a concept that remains hard to define objectively. Therefore, I have chosen a definition that contains some subjective elements (well-run club as an example) in order to reflect the nature of the concept. Since competitive balance influences the demand for the sport, the perception of its existence is more important than the proof that it really exists. No one outside of academic circles will sit down and plug the numbers in a formula to see if

---

118 This is explored in details in Chapter 3.
competitive balance exists. As reported by Mr. Costa, fans will recognise it when they see it. For this reason, I think that this definition is the ideal one.

Since competitive balance influences the revenues of the team, it is an important factor in the market for players. If the teams make more money, there are good reasons to believe that the players’ salaries will also grow.

With that in mind, I will examine in the next chapter how value is determined in professional sports leagues, with the emphasis on the value of teams and players.
Chapter 3 - The determination of value in professional team sports

“Owning the Yankees is like owning the Mona Lisa.”

- Attributed to Georges Steinbrenner, owner of the MLB New York Yankees

With this small sentence, Mr. Steinbrenner expressed the difficulty of trying to value anything objectively in professional sports. One of the main problems is the lack of information that makes it difficult for an outsider to make more than an estimate. Professional sports teams are usually private entities, and even if they are public corporations that does not guarantee more information on their inner workings. The public corporation form only partially solves the information problem for outsiders since there are ways to keep information vague even when the securities regulations force a certain level of disclosure. Such an attitude can be partly blamed on the culture of secrecy that is prevalent in the professional sport industry that could be attributed to the ultra-competitive nature of the industry: every bit of information is guarded carefully in the hope that this knowledge will result in some kind of advantage for the team.

This chapter will approach the determination of value in professional team sports in two steps, first I will look at the teams and the way that their value is determined and after that, I will look at the value of players.

I will approach the value of teams by examining some of the corporate structures that they use to conduct their business and how it differs from the way a normal business would conduct itself in the same situation. In order to understand that difference, I will have a look at the theory of the sportsman owner to give a better understanding of the incentives that motivate the owners of sports teams. Looking at the value of teams is important since it influences the way the players are valued, the player being the most expensive asset of a team.

For the value of player, I will compare two ways to determine the value of players, looking at it either from the revenue function perspective of the team or from the economic surplus of team perspective. In the final part of the analysis I will make a decision about which definition that is going to be used for the remaining of the thesis.
3.1. The situation as it stands now for determining the value

The best estimates for North American professional sports teams are done every year by the financial magazine Forbes. Usually teams and leagues disagree with Forbes’ final numbers, especially when they go against the ones they made publically in order to get public subsidies, usually to build new stadiums.

In European Soccer, the accounting firms Ernst & Young and Deloitte publishes yearly data about the sport, but that are again only estimates based on incomplete data and educated guesses.

As for the determination of the value of players, the question is not one of lack of information since most of their performances and salaries are public and subject to numerous analysis. Here the difficulty is to find a standard way to put a value on these performances.

3.2. The value of teams

Trying to determine the value of professional sports teams places us in front of three problems.

The first obstacle is, as mentioned in the introduction, the culture of secrecy that surrounds the finances of professional sports teams. They seem to be afraid of the reaction of their fans and customers if they reveal too many details about their commercial interests. Fans would be able to see that price increases are mainly due to bad management and not to an investment in better players. In addition, teams are the beneficiaries of numerous tax exemptions, legislative advantage and subsidies. All that could be endangered if the public knew the real extend of their finances.

The second obstacle results from the flexibility of the generally accepted accounting principles. This flexibility allows every business to “play around” with their financial reports - in the case of a professional sports team the effect can be extreme. As Paul Beeston, owner of the MLB’s Toronto Blue Jays once said
“Anyone who quotes the profits of a baseball club is missing the point. Under generally accepted principles, I can turn a $4 million profit into a $2 million loss and I can get every national accounting firm to agree with me.”

This comment could apply to other industries, but in professional sport, the secrecy, the integration of teams in broadcasting holdings and the difficulty of establishing an objective value of the assets of the teams seem to compound the problem.

The third obstacle is the difficulty of establishing the value of players. Players are the most important assets of a team and, as we will see later in this chapter, they are hard to value correctly. With that much uncertainty regarding their main assets, the accounting value of a professional sports team can vary depending on the wishes of the owners and their accountants. But accounting value as only a limited value for the purpose of this work, the real debate is to see if there is real economic value in the teams. One of the rare occasions that allow us to look inside the financial situation of teams and leagues and their values is when there is a change of ownership.

3.2.1. The market value of teams

The sale of Manchester United to the American Malcolm Glazer for 790 million £ in 2005 can give an indication about the value given to English soccer clubs by eventual buyers. However, the declaration made by Mr. Glazer in 2007 saying that he makes more money with his relatively unknown NFL team in Tampa Bay than with a world renowned soccer club allows us to question the sanity of this purchase or the sincerity of the declaration. Why pay that much money for a team that is, according to the owner, not worth it?

Another example is the takeover of the FC Liverpool by two wealthy Americans, George Gillett and Tom Hicks. The net value of the transaction was estimated at 174.1 million £.

---

In addition, the two new owners made investment promises for a new stadium and training facilities that should boost up the cost of their investment up to 470 million £\textsuperscript{124}.

An interesting side note to the Liverpool story may shed some light on the incentive of some prospective football club owners and help us understand some of their motivations. Messrs. Gillett and Hicks were competing with another bidder for the club; Dubai Investment Capital (DIC), a company owned by the ruler of Dubai. As the two groups fought for control of the FC Liverpool, an interesting document prepared by DIC was leaked to the BBC. This document explained the reasons why DIC wanted to invest in the club. It indicated that DIC saw FC Liverpool as an investment like any other and explicitly stated that it wanted to develop the commercial business of the team and expected to make a return on capital of around 25% a year, based on similar acquisitions made in the English soccer market. Additionally, it stated DIC’s intention of selling the club after seven or eight years and cash in on their investment.

This document shocked the managers of FC Liverpool who apparently did not appreciate being seen as a pure commercial venture and they refused the offer of DIC\textsuperscript{125}.

That DIC would expect such a return on their investment should not be surprising. The financial situation of the English Premier League has never looked better. Since its creation in 1992, the revenues of the League have increased tenfold; that is a 1000% increase in revenue in 15 years\textsuperscript{126}. Before the 2007-2008 season, the League signed a new television broadcasting contract that brings it in 3.5 billion £ per year, the highest television revenue of any soccer league\textsuperscript{127}. This growth in revenues helps explain the high price, but also the surprising liquidity of these clubs in a market with relatively little supply. Before the sale of Liverpool and MU, Chelsea was bought by Russian billionaire Roman Abramovich for 123 million £\textsuperscript{128}


\textsuperscript{127} Chapter 2, section 2.2.1.

\textsuperscript{128} MAHONEY, R., 2007. English soccer fans wary as Yanks take over teams. USA Today, No. 4758, May 16, p.1A.
and the American Randy Lerner bough Aston Villa for 65 million £\textsuperscript{129}. Maybe one should question Mr. Glaser’s sincerity when he complained about the low revenues he was getting out of MU. The structure of the English clubs that are mostly publicly traded companies or private corporations makes the acquisition of the club relatively trouble-free for outside investor. But the market value of teams does not make for a good measure of the objective value of teams since the conditions of each sport and each market are very different and make comparison difficult.

In addition, since the teams are owned using different ownership structures in different sport and market, that makes the comparison even more difficult. To illustrate that problem, I will now look at some examples of ownership structures and their influence on the value of teams.

3.2.2. The publicly traded corporation in sports

The publicly traded corporation is a strange form of organisation for sports teams. Although there are financing and fiscal advantages, it is hard to see the incentives for the investors to put their money in a sports team. These kinds of companies have been underperformers if compared with all other industries over a long period and traded at prices not justified by traditional stock analysis\textsuperscript{130}. This would mean that the investors, when buying a share of a team, either act irrationally or they must derive utility from the pure possession of a share of the club, utility that compensates for the lack of traditional value of the stock. This theory seems very credible since the fears of sports teams’ shares owners during ownership changes are different from the ones that characterise traditional takeovers. The debate that raged during the acquisition of Manchester United by the American tycoon Malcolm Glazer proves at least part of this. The individual small stock owners, mostly fans, owning fewer than 2% of the shares, were not complaining about the price that they were getting for their shares. Their main complaint was that Manchester United tickets and merchandising prices would be raised; that the club would be loaded with debts and that financial details would not be public anymore\textsuperscript{131}. These are not the typical shareholders’ concerns in a takeover.

\textsuperscript{129} Idem.


3.2.3. US teams, publicly traded corporation and their value

Past experiences in the USA showed that it is hard to find rationality in the investment of small shareholders in sports teams.

In an attempt to raise money, the NFL Green Bay Packers issued 400 000 shares in 1950. On the opening page of the prospectus one could read

"It is virtually impossible for anyone to realize a profit on a purchase of common stock or even to recoup the amount initially paid to acquire such common stock".\textsuperscript{132}

A study\textsuperscript{133} conducted about stock sales in the four most important professional sports leagues in the USA showed that these sales are seen by the fans as a means of showing support for their favourite team rather than a financial investment. The investors are satisfied with framing their stock certificate with no hope of making any money. Of the four teams examined in the study, only one actually paid any dividends to the owners of the shares. The other teams specifically wrote in their prospectuses that no dividends would be paid. The selling of stocks by professional sports teams in the USA seems to be a revenue generation scheme and not a genuine equity financing structure. These kinds of shares allow the teams to get the benefit of fresh money without the strings that are usually attached to equity financing\textsuperscript{134}. In addition, since all US leagues reserve themselves the right to veto any change in ownership of teams, it strips these shares of their remaining value: it makes them illiquid and almost impossible to sell. Since the shares are now striped from two of the three legal aspects of a share, their value is greatly reduced. Without the right to a dividend and ownership rights there is little interest for a potential buyer to attach any value at all to this share.

So if minority ownership of shares in a sports team is only a way of supporting his favourite team, is majority ownership of a team any different? Is there a way to explain the actions of the owners of sports teams that seem to lose all their business acumen when managing their

\textsuperscript{134} Idem.
team? Do they act rationally and is there a way of explaining their behaviour? A possible explanation is called the sportsman owner.

3.3. The sportsman owner

This model has been developed for North American sports\footnote{VROOMAN, J. 1997. A unified theory of capital and labor markets in major league baseball, 63 Southern Economic Journal, 594 – 619 and VROOMAN, J., 2000. The Economics of American Sports Leagues, Scottish Journal of Political Economy, pages 364–398, September 2000.}, but it also applies to the European owners. It is the result of the observations that sometimes sports teams’ owners do not act the way that could be expected from a normal profit maximising owner, but they do not act as a purely win maximizing owner either. The sportsman owner offers a third way that could better help explain the action of the owners of sports teams. The model first makes the assumption that the traditional pure profit American sport owner model is wrong. Instead, we face a “sportsman” owner who is ready to trade off some profits for the satisfaction of winning games. This is what is called the “sportsman effect” and means that in part, profits in the utility curve of the owner are replaced by wins. The ratio seems to be different for each owner, but the more valuable a win, the stronger the effect. Since profits are not the sole motivation of the owner, it has consequences on the value of the teams.

A league with sportsmen owners will always have a lower average team value than with the traditional pure profit model. Winning being a zero sum game (you either win or you lose) it is impossible for all sportsmen owners to compensate their losses of profit with wins. So even if all owners try to have a better team they cannot all reach their objective since the number of available wins in a season will remain the same. So in order to get more wins, the sportsman owner will try to get the best players for his team. This will also result in a loss of team value from the higher players’ salaries. The more sportsmen owners the league has, the stronger the loss of value will be. Sportsmen owners are willing to pay more in order to have better on field results. The average salary will subsequently increase since there is only a limited supply of good players and all the teams will compete to get them on their squad. This increase in players’ salary will result in lower profits for the teams since they will have higher expenses. These lower profits result in lower financial value for the team when compared with a pure revenue maximizing team. So the presence of the sportsmen owners results in a decrease of the value of all the teams in the league as each one must compete for the best players.
But this is not the only effect of the sportsman owner on the value and structure of professional sports. The motivation of these owners leads to the use of a new legal form that also reduced the value of teams, the “leveraged syndicate”.

3.3.1. The leveraged syndicate

The use of the leveraged syndicate as a legal entity in professional sports was popularised by Georges Steinbrenner, the former owner of the MLB New York Yankees. Leveraged syndicates are basically partnerships with the different partners being given shares in the partnership. There are two types of shares: general partnership shares and limited partnership shares. Owners of general partnership shares, called general partners (GP), are given the right to manage the corporation and they are usually not required to invest any money in it. Managing the corporation means that a GP can make all the business and day to day decision that he deems necessary for the good conduct of the operations of the firm. But the GP is ultimately liable for the debts of the corporation; his liability is unlimited just like in the case of a normal partnership.

The owners of the second type of shares, called limited partners (LP), are only investors who give money and allow the GP to do what he wants with the investment. Their shares do not give them any right to manage the corporation and they must remain passive investors. In return, they will receive part of the profits made by the syndicate. Should the corporation be forced into bankruptcy, the liability of a LP will be limited to the value of his initial investment made in the firm. However, in order to benefit from that protection, the LP must not act as or do actions that are deemed to be made by a manager, he must remain completely passive. That means that the GP basically has the freedom to do whatever he pleases with other people’s money, within the constraint of the partnership agreement. This situation gave rise to the most interesting quote from one of Mr. Steinbrenner’s LP’s “There is nothing more limited than a limited partner in the Yankees.”

3.3.2. The sportsman owner and the leveraged syndicate

A sportsman owner using a leverage syndicate has an even greater negative effect on the value of the teams in a league than one using a standard partnership or a publicly traded

---

136 Attributed to John McMullen, former LP of the Yankees.
corporation. This comes from an imbalance of incentives resulting from the structure of the leveraged syndicate allowing one party to manage the team using the money provided by another. The GP has all the benefit of a winning team, all the utility of the wins, but only part of the negative impact coming from the lowered value of the team because of the lower profits and revenues resulting from the investments, mainly in players, required to achieve these wins. To make matter simple, one can say that he only suffers 1/N of the loss of value, N representing the number of limited partners in the leveraged syndicate, but gets all (1) the benefits that come with winning. The GP therefore has a greater incentive to hire the best and most expensive players since he is the sole recipient of the utility of the wins but pays only part of the costs; the situation gives him all the right incentives to over invest in the salaries of players. In short, GPs play with other people’s money and have little incentives to make any profit in order to repay the LPs. The latter end up being the fools in that situation since they pay part of the costs and get no return or benefits of winning as they are forbidden to take part in the management of the team. The lowering of the team value due to the combination of the sportsman owner and the leverage syndicate can also be aggravated by the negative influence they can have on competitive balance. As salaries get higher, only a limited number of teams, the richest ones, can keep up with the salary inflation. Therefore the chances of having a balanced league get lower and lower as the post-season places get monopolised by the same rich teams year after year. Having an unbalanced league should also lead to lower revenues since unbalance will result in lower demand for the product of the league137 and lower the value of the teams even more.

3.3.3. The empirical problem with the sportsman owner

However, there is an important paradox between that theory and the empirical facts in North America. Although most teams are organised in the form of leveraged syndicates, their value keeps growing. In 1990s, there was a rate of appreciation of around 15% for the MLB teams138 in spite of the fact that MLB revenues had only grown by 11% for the same period. Does that mean that the model is wrong and that the sportsman owner does not really exist? A plausible explanation is that the discordances between the models and empirical results come from the unchecked monopolistic power of the MLB owners that allows them to extract extra

137 See Chapter 2.
138 SCULLY, G., see note 101.
revenues and monopolistic rent. That would mean that MLB, and other sports team owners, only get a return on their investment because of the absence of competition in the market for sports leagues.

Another explanation is that the basic assumption of the sportsman owner concept -saying that owners are not purely profit driven- is incomplete. The sportsman owner model does not account for the demand side of the equation of the sport business. The revenues of most professional sports are growing very rapidly since the last decades of the XXth century. If the revenues grow faster than the decline of value resulting from the growing salary bills, this could explain that the value of teams seems to be growing all the time and would not mean that the owners disregard profits.

Yet another point to mention is that there are problems with the empirical measures. The value of teams and their growth rates are usually established using the market value of teams. But the market for teams is very illiquid and the rationality of investing in them is not well understood. As seen earlier, stocks and investments in sports teams do not seem to respond to the same rules as other investments. Owners seem to get utility from the simple fact that they own a sports team. This has been called the ego premium and represents the ego gratification coming from the ownership of the team and being associated with it. In addition, the reasons of acquiring a sports team may have nothing to do with the revenue function of the team and more with its combination with other assets that the owner has. Especially in the case of media industry, the combination with a sports team has been considered equivalent to vertical integration or cross subsidization. In these situations, the owner may want to maximise the revenues of the media outlets by giving them the rights to broadcast the games of the team at a below market rate. The costs saved by the broadcaster could be interpreted as a subsidy from the sports team and contribute to lower the value of the team. The contrary could also be true, the owner forcing the media outlet to take losses by selling the rights to broadcast the games above the market price.

---

139 VROOMAN, J. 1997, see note 135 and VROOMAN, J., 2000, see note 135.
141 See section 3.2.2.1.
143 Idem.
3.3.4. European soccer and the sportsman owner

The traditional European model states that managers behave like leverages sportsmen owners. They have an incentive to overinvest in players’ salaries since they try to maximise their own individual utility through the success of the team. Like in a leverage syndicate, the manager does not play with his own money but is the main beneficiary of the fame and glamour that comes with a winning team so it gives him the incentive to invest more than the equilibrium in playing talent and overinvest in players’ salaries\textsuperscript{144}.

3.4. What the value of sports teams really is about

The value of a sports club is not derived from the players’ contracts or from the assets that he owns. What determines the value of a team is his part of the monopolistic rent that comes with his membership in a league\textsuperscript{145}. The value of the rent does not usually diminish over time unless a rival team of the same league is established in the exclusive territory. In closed leagues, teams have a veto on the establishment of new teams. That allows them to protect their monopolistic rent and therefore their value. This is not an option for teams in open leagues since they have no control on the localisation and the identities of the teams that enter or leave the league. There is therefore no territorial monopoly in these types of leagues but a participation monopoly coming from the fact that the number of places in the league is limited. Even if the league is open that does not mean that there is no monopolistic rent to exploit. By limiting the number of clubs that can take part in the first division competition and the number that will be relegated/promoted, soccer organisations create a monopolistic rent for the clubs allowed to participate although that rent does not include a territorial monopoly. This lowers the value of a team compared with a similar one in a closed league, but does not eliminate the monopolistic rent that the teams are able to extract from the market.

In conclusion to that section, we may say that the value of a professional sports team mainly results from its exploitation of its part of the monopolistic rent that comes with its membership in the league. This is the basic business model that allows teams to charge the price that they want for tickets, merchandising and broadcasting rights. It is from that exploitation of the monopoly that they get their revenues with which they pay their players\textsuperscript{144, 145}.

\textsuperscript{144} DIETL, H., FRANCK, E. AND LANG, M., 2005, see note 81.
\textsuperscript{145} LEEDS, M. AND VON ALLMEN, see note 142 and ZIMBALIST, A., 2001, see note 140.
and which is irrespective of the motivations of the owners or the ownership structure of the team.

3.5. Determining the value of professional sport players

The purpose of this section is to try to establish an acceptable way of determining the value of a player in professional team sport. In order to find something acceptable in the context of this thesis, we will first have to examine the traditional approaches of determining the value of a player and see if they fit the purpose of this work. At the end of this section, I will establish a way of determining the value of a player that will be used for the remaining of this work.

3.5.1. Determining the player’s value with the revenue function of the team

One can measure the value of a player by trying to determine the marginal revenue product (MRP) function of the player. This method assumes that the team’s on-field performance is a function of the players’ production and that a team revenue function has a very close correlation with its winning percentage. The goal of this approach is to estimate the individual contribution of a player to the team on-field performance. In this case, the MRP will represent the individual contribution of the player to the on-field performance of the whole team. Using the player’s individual statistics, one can estimate the dollar value of his MRP\(^{146}\) creating a ratio between the on-field performances to the revenues that are received by the team as a result of these performances. This method allows for a clean and mathematically satisfying calculation of the player’s value. In the end, it gives us a nice model that is proved and solid. But it still remains limited by the inputs that are used in the model. We therefore need reliable individual statistics that allow us to see how the player influences the chances of winning of his team. The problem with this is that some facts are not reflected by the statistics: defensive ability is hard to put in number for individual players, leadership and tactical intelligence are impossible to quantify. Even if we assume that we are able to use statistics to reflect all the skills of a player, would that accurately reflect the reality? Does the MRP reflect the way teams think about their players?

---

The main problem that we face here is that by correlating wins with revenues, this model assumes that the sport business is a zero sum game in which the winner takes all. By assuming that teams with a winning record will get more revenues, it evacuates competitive balance which is a mistake since it is probably the most important factor in the revenue function of a team as we have seen in chapter 2. To use the vocabulary of this chapter: the model assumes that all the consumers are fans of the team and not fans of the game. The main reason for this mistake is that this model was developed at a time when professional sports got most of their revenues from stadium attendance. In today’s world a growing portion of the sports teams’ revenues comes from broadcasting multi-year contracts less affected by the teams’ annual performances. Therefore to perfectly correlate wins with revenues does not reflect reality.

In addition, the model completely excludes the commercial off-the-field value of a player so that for example a very popular player will not be considered more valuable for the team if this popularity is not justified by his on-field performances.

Therefore, this is not the correct method to use for this thesis.

3.5.2. The willingness to pay for the services of a player

The usual way in North America to determine the team’s willingness to pay is to use the player’s salary, i.e. the amount of money the team pays him. These numbers are usually published by leagues in accordance with a section of their collective bargaining agreement to allow players to get more information about what their colleagues are paid and allow them to have better information on the market for their services. For European sports, the amount of the transfer fee paid in order to acquire the player needs to be added to his base salary and the total sum is considered the value of the player. This way of valuing a player reflects the willingness of teams to pay for his service.

Willingness to pay is commonly used in economics as a way of finding the value of things, including human resources. There are no good reasons not to apply this method to the market for players and it has the advantage of being relatively easy to understand once one has acquired the information needed. In order to perfectly reflect the value of players with that

147 See chapter 2, section 2.2.1.
method, we have to assume that all the parties involved are perfectly informed regarding the characteristics of the players. But this is far from being true. In reality, information asymmetry results in cases where players seem over or underpaid relative to their on-field performances. Still, teams are the decision makers in the market and by using this demand oriented method, we at least reflect the conclusions they came to when hiring players. This is why examining how to determine the value of a team is so important.

If the market for the service of professional sport players was a perfectly competitive one, the salary of the players would be an accurate measure of their values. But for this, we need to assume that each party has perfect information and that the performance of the player can be measured objectively. Fulfilling these assumptions is very difficult in reality as already observed earlier. Measuring the performance of a player is a perilous operation. Some facts like goals scored are relatively simple to observe and record, but the defensive contribution of a player stays unrecorded and is usually evaluated subjectively: a tackle to stop a dangerous attack will never get recorded but a goal in a meaningless 7-1 defeat will. The objective of the team in determining their willingness to pay, like in the MRP method, remains to correctly price the value of the player according to his contribution to the team. One should be able to isolate his performance from the rest of the team and see how valuable the player is. Part of the problem is that the evaluation of a player can be highly subjective and his value may change depending on both the observers and the team the player plays on. For the purpose of this work, willingness to pay of teams is the best way to measure the value of players. We do not need to a sure way of measuring objectively the on-field performance of players, since it is only part of the value of the player.

Willingness to pay affects some degree of flexibility and allows for subjective factors to be taken into consideration. Willingness to pay may be one of the best methods to determine the value of a player because it does not have the problem of trying to fit into objective data. By allowing for the subjective judgment of teams and players, right or wrong, this leads to a realistic evaluation of the value of players.

3.5.3. The surplus of teams and its influence on the willingness to pay for players

If we assume that the willingness to pay for players by teams constitutes the most realistic way of representing the value of players, it is necessary to try to examine how teams
determine their willingness to pay for a player. In order to do so, teams will try to calculate the value of the surplus that is going to remain in their coffers at the end of the transaction with the player. The first question that one must answer is: what are the costs related to a player? The costs associated to the player are: his salary, the transfer fee paid to acquire him or the training costs to develop the player. For the purpose of this exercise, I will merge the transfer fees and the development fees related to the player in one category called the acquisition costs. This category will represent the non-remuneration aspect of the costs related to the player. These costs are relatively easy for teams to calculate since they have direct access to this information.

So assuming that the teams are rational, their objective will be, for every player, to end up with a positive Economic Surplus (ES) or, at the worst, a value of 0 for the following equation:

\[ ES = Fv + Cv - S - AC \]

ES: Economics Surplus  
Fv: Field Value  
Cv: Commercial Value  
S: Salary  
AC: Acquisition costs (transfer fee and/or training and development costs)

In a perfectly informed and competitive market, ES would always be 0 since the player would be paid at his correct value and S could be instantly adjusted to reflect the fluctuating Fv and Cv. To determine Fv the team will ask itself how their on-field performances would change if they added/subtracted the player from their squad. Depending on the team, it could be a relatively sophisticated or quite simple calculation. Cv is an often neglected aspect of the valuation of a player. It can be represented as the non-playing parts of the equation: the sponsorship agreements, the jerseys and other merchandising objects, etc. Merchandising directly linked to the player is relatively easy to measure, but it is more difficult to find an individual player’s contribution to the general commercial revenues of the team. A simple way to measure that would be to represent the contribution in the following way: \( \frac{1}{N} \times \text{percentage of the season played by the player} \times \text{the total general commercial revenues of the team} \). N is the number of players on the team’s roster. It is a crude method that however
reflects reality rather well. Only in exceptional case of superstars does the commercial contribution of a player require a more complicated method\textsuperscript{148}. But superstars are a minority and for our purpose, one could control that by giving them manifold times the importance of a single player (it may be imperfect but compared with the MRP method, at least this equation accounts for this superstars phenomena).

3.5.4. Multiple years contracts

Another factor to consider in relation with the players is that they usually sign multi-years employment contracts. In order to conclude these contracts, teams have to make a prediction about the development of the value of the player in the future. It may grow or diminish, for reasons both external to the team/player relationship like an injury or a national team selection and internal such as the development of a better young player playing at the same position and taking away playing time from the player. This makes predicting the future value of the players even harder. And the same thing can happen regarding the commercial value of a player, creating even more problems for the team when trying to determine the long term value of the player.

This is probably the source of most problems when clubs try to determine their willingness to pay. Mistakes made in a multi-years contract on any of the variables will compound themselves and their effect will become stronger the longer the duration of the contract.

3.6. The definition of value of a player for the purpose of this work

Despite the problems mentioned earlier, willingness to pay seems to be the best method of determining the value of the player. This shall be the basis upon which I will construct the definition of value of players for the purpose of this thesis. It represents an accurate portrait of the reasoning and incentives that teams face when trying to determine the value of a player in the context of changing regulations.

But the weaknesses should not be forgotten, teams and the players themselves can find it difficult to come to a correct evaluation of the on-field value of a player over the course of a multi-year contractual relationship. The pure statistics do not always perfectly reflect this

value and the traditional way of compiling them may be totally inappropriate and inaccurate to provide a correct measure of the player's contribution or the statistics that are available can be used in a suboptimal way.\textsuperscript{149}

In addition, managers, players and team personal tend to overvalue the on-field value and undervalue the commercial off-the-field value of a player. The more players are on the field, the less important the individual contribution of a player is. One star player cannot bring a soccer team to the championship all alone. In addition, a playing season has a limited number of games and the on-field value is not worth anything during the off season. It is also under the influence of external factors the worst of them being injury. All this drops the FV to 0 despite all the amount of care of both parties in the contractual relationship.

The commercial off-the-field value of a player is, however directly, attached to him. It is not entirely dependent of the success of the team; the commercial value is not linked to the season or reduced to 0 by an injury. It is a safer investment as it tends not to be too strongly correlated with the performances of the player. Once the player becomes a pop icon, his commercial value is safe from any bad performance on the field and he can still cash in even when retired. But the perception of the parties implicated in the business is that it is of the same nature as the on-field value, a good of finite nature. To quote Mr. Julien Brisebois of the Montréal Canadien “What do we do with a player that has been with the team for a couple of years? His commercial value is not worth that much since he probably sold all the jerseys that he could sell”.\textsuperscript{150} This may be accurate, but one could do the second best thing which is in fact standard practice in Europe soccer: change the uniform of the team.

So keeping both these limitations in mind, for the purpose of this thesis, the term value of a player will mean the amount of his salary and the acquisition fee paid by the team to acquire him. By acquisition fee I understand both the costs of training and developing the player.

Now that we have a satisfactory definition of the value of a player, we can look at how this value is influenced by the changes in the various rules of professional sports. In the next chapter, I will address the changes in the rules of the game and their influence on the value of players.

\textsuperscript{150} Interview with J. Brisebois, see note 114.
Chapter 4 – The effect of a change in the rules of the game on the value of players

- The point of soccer is to equalise before the other team scores.
  - Robert Dennis “Danny” Blanchflower

4.1. The rules of the game

Sit on a bench near a playing ground and look at a group of children playing. The first thing that they will do, after having agreed what game to play, is to agree on a set of rules that are going to regulate their game. This rule instinct probably comes from a natural sense of competition, a desire to find out who is the best. In order to do so, we need a levelled playing field. This is a corollary or a consequence of the Lewis-Schmeling paradox in which the rules help to make the contest fair for all competitors.

Disagreements about the rules of a game are at the basis of some of the most important events in the history of sports. Disagreement about what a tackle was resulted in the creation of two different sports. But as far as we can tell, the rules of the game in all professional sports are now crystallized in codes and rule books that are as hard to amend as some countries’ constitution.

Such conservatism and stability is not a bad thing per se. It allows the players to develop a set of skills, invest in their development and plan a long term investment in abilities and skills in function of these rules without the fear that a rule change might suddenly devaluate them or make them obsolete. It also allows the teams to prepare themselves and their strategies knowing what is allowed. Without such stability, teams would be forced to plan from one year to the next losing the advantages provided by long term planning and investments.

The goal of this chapter is to show the influence of the rules of sports, the rules of the game on the field, on the value of players. In order to do so, I will mainly examine changes that were made in the rules of the game and look at the influence these changes had on the value of players. By using this before and after approach, I will be able to actually see the difference in the valuation of players. The main problem with this is that not all sports changed their rules or changed them in a way that allows measuring a global influence on the value of players.
players. It may affect only a certain class of players and therefore be of limited use. It is hard to keep a comparative approach in such a situation where some sports like soccer basically still play with the same set of rules as when they started whereas others annually change some aspects of their rules of the game. In order to draw some useful conclusions, I will try to present the changes in the rules of the game that are the best examples and that can be applied across all sports. I will look at a change in the rules of the game of one sport, the effect it had on the value of the players in that sport in particular and in the end provide an explanation why this can be generalised to all sports. The goal is to find a common denominator, a common factor that can be applied to all sports. This will be very useful since it allows us to find out what change of the rules of the game will influence the value of players without actually having to change the rules. That way of compensating for the lack of data should at least allow us a satisfactory generalisation of the effect of changes in the rules of the game on players and make the impact of such change clearer.

4.1.1. Definition of the rules of the game

What are the rules of the game? In order to examine them, we first need to actually find a definition of that aspect of sports. Trying to do so, we are faced with the problem that each sport seems to have a different definition of the rules of the game. This divergence has its origin in the legal cases that were fought by the various leagues over the years. In most of these cases, if the rule contested did not regulate an economic activity, then the courts usually did not intervene as we will see later in this section. Therefore, it was in the interest of sport organizations to try to label all rules of the game as non-commercial, with some unforeseen consequences in the end, especially in Europe, as we will also see in this section. In order to show that evolution, let us examine some important ECJ rulings on the applicability of the Treaties on the rules of sport. This analysis will allow us to see how the ECJ vision of the rules of sport evolved over time.

The first case to examine the applicability of Community law to sports was the Walrave vs. Union Cycliste Internationale\(^{151}\) in which the Court stated that the practice of sports is subject to Community Law only as far as it constitutes an economic activity within the meaning of Article 2 of the Treaty. That opened the door to European judicial intervention in the world of sports. But the ECJ was quick to impose limitations as it can been seen in the Donà vs.

Montero\textsuperscript{152} decision in which the ECJ stated in what must be considered as an *obiter dictum* since it was not necessary to adjudge on the case that involved professional and semi-professional soccer players and had nothing do to with national teams

>“However, these provisions do not prevent the adoption of rules or of a practice excluding foreign players from participation in certain matches for reasons which are not of an economic nature, which relate to the particular nature and context of such matches and are thus of sporting interest only, such as, for example, matches between national teams from different countries.”\textsuperscript{153}

For almost 30 years, that view had been confirmed up to various degrees by the ECJ. In *Bosman*\textsuperscript{154} and other decisions like *Deliege*\textsuperscript{155}, the ECJ confirmed that the national team selection rules and the composition of the national team itself decided by a sport national association were excluded from the application of the Treaties since these did not constitute economic aspects of sports.

In 1999, the Commission issued a report\textsuperscript{156} in which it tried to make sense of the various ECJ sports related decisions. According to the Commission, the “rules of the game” are the ground rules, the rules of the National Associations and the rules necessary for the organization of competition. They include all rules that have for sole purpose to regulate the sporting aspects of the game. That includes, notably, the size and the marking of the field, the number of players per team, the legal playing actions available to a player, etc. The goal of these rules is not to distort competition in sports but to allow it. Therefore, they should be considered as non-economic sports rules. The legislator did not intend to regulate these rules and the leagues and associations are rather free to do what they want concerning that aspect.

However, this view was partly contradicted by the ECJ itself in the *Meca-Medina* decision\textsuperscript{157}.

After stressing that not only the non-economic aspects of a rule should be considered but also the objective and effect of such rule closely examined the ECJ wrote

\textsuperscript{152} [1976] ECR 1333.
\textsuperscript{153} Idem.
\textsuperscript{154} See chapter 5 for a detailed analysis of the Bosman decision.
\textsuperscript{155} Christelle Deliege vs. Asbl Ligue francophone de judo and others, C-51/96 and C-191/97 (joined).
\textsuperscript{157} Meca-Medina and Majcen vs. Commission, C-519/04.
“(…) it is apparent that the mere fact that a rule is purely sporting in nature does not have the effect of removing from the scope of the Treaty the person engaging in the activity governed by that rule or the body which has laid it down.

If the sporting activity in question falls within the scope of the Treaty, the conditions for engaging in it are then subject to all the obligations which result from the various provisions of the Treaty. It follows that the rules which govern that activity must satisfy the requirements of those provisions, which, in particular, seek to ensure freedom of movement for workers, freedom of establishment, freedom to provide services, or competition.”

One must remember that the Court was examining in that case the effect of the anti-doping rules of a sport on the free movement of person. What the ECJ did in that decision was make all the rules of a sport, per default, subject to Community law. It is possible to resume the ratio of the ECJ in Meca-Medina in a simple two steps test. First, one must consider if the sporting activity in question falls within the scope of the treaty. If the answer is yes, then one must ask if the activity currently examined can be regarded as an economic activity. Before this decision, it seemed to be the other way around. This may look like a negligible detail but it is not. With the Meca-Medina decision, all the rules of a sport are per default considered subject to community law; there is no need to first show economic activity requirement. In short, the burden of proof has been reversed from the party asking for the application of Community law to the sport authorities wanting an exemption.

For the purpose of this thesis, the rules of the game will be defined as the non-economic rules of a sport. However, as the Meca-Medina decision shows, it is actually hard to find a rule that does not affect, one way or another, the economic aspect of sports. In some ways, for example the number of players on the field and the duration of the game could be considered to have an impact on the economic aspect of sports. In the case of the number of players, it determines the size of the work force and the wage costs and for the duration of the game, it may have an impact on the value of broadcasting contracts. So, in order to refine the definition even more, the rules of the game, for the benefit of this thesis, should be understood as rules, adopted by the various sport organizations, which have no direct economic impact on the industry; notably, the rules in which the sport is conducted on the field of play and the national team selection rules. This evaluation has to be done on a case by case basis, but is compatible with the ECJ jurisprudence mentioned above.
4.2. Changes in the rules of the game

There are many ways a sporting organisation may change the rules of the game, but for the purpose of this thesis, I will put them in two broad categories that easily include all possible rules changes: radical changes to the rules of the game and mild changes to the rules of the game. The smaller the change, the harder it is to measure how it influences the performance of the players.¹⁵⁸ I will first examine a radical change to the rules of the game and its effects on the value of players and in the following section, I will look at the effect of a mild rules change. One of the main problems that I encountered trying to make generalisations for the rule changes is, as said earlier, that it is hard to find sports in which rule changes actually occurred. Rule committees are usually very conservative and reluctant to change the established tradition of sports, so the occasions to actually see a change in the rules of the game are pretty rare. There are, however, two North American examples that, from their nature and from the impact that they had on their respective sports, allow us to make some interesting conclusions on the impact of rule changes in a general sense. For the radical rule changes, I will examine the introduction of the forward pass in football and for mild changes the lowering of the pitcher’s mound in baseball. The following section will allow the reader to understand why I chose these examples and why they allow us to make generalisations that apply to all sports.

4.2.1. Radical changes to the rules of the game

As said in the previous section, studying the radical changes in the rules of the game allows us to have a clearer indication of their impact on the value of players. The main goal of studying these changes for the purpose of this thesis is to observe the changes in the value of players in extreme cases. This should help develop a measuring standard for the more subtle cases. It also provides the ultimate test for our hypothesis that rule changes influence the value of players. If this does not happen in the most extreme cases, then other types of rule changes also should not matter.

What I mean by the label radical changes in the rules of the game are the changes that greatly affect one aspect of the game or introduce a new one to the sport. Such changes are easy to imagine, like allowing soccer players to use their hands, but hard to find in the modern history

of professional sports. In order to find one such example with enough data to make any kind of useful conclusion, I had to go back to the beginning of the 20th century.

4.2.1.1. The introduction of the forward pass in football

At the beginning of the 20th century, football was a very dangerous sport. With the rules in force at the time, there were little tactics and strategies involved. Games mostly consisted of masses of players banging on each other to make the ball progress, much like the scrum technique in Rugby. In 1905 no less than 19 players died on the field. This situation allegedly compelled USA President Theodore Roosevelt to summon the presidents of the three most important football playing Universities - Harvard, Yale and Princeton - to discuss the state of the game. Legend has it, since there are no written records of that meeting, that Roosevelt delivered an ultimatum to the universities: if they did nothing to quell the violence of the game, he would ban it. At the start of the 1906 season, the universities football rule committees introduced a rule that allowed one player to make one forward progressing pass to one of his teammate. Such an action was originally, like in modern Rugby, forbidden. According to the universities, this new rule was adopted:

“(…) with the double object of eliminating brutality and making the game more interesting for the spectators.”

The goals of the football authorities were clear but were they actually reached?

The first goal, to reduce the amount of deaths and injuries on the field, was certainly achieved. Two years after the introduction of the rule, there was a marked decrease in death and serious injuries suffered by college players. In the 1905 season, in all the USA, 19 college and high school players died; in 1906 11 players suffered the same fate and the same number (11) in 1907. In other football leagues that did not introduce the forward pass, the amount of death and injuries stayed constant. Although a two years period is a little short to draw radical conclusions, one can nevertheless say that the introduction of the forward pass at

---

159 NEW YORK TIMES, 1907. Football’s Death Record for 1907. November 24.
162 NEW YORK TIMES, 1907, see note 159.
163 Idem.
least contributed to the reduction of the “brutality” of the game; deaths on the field are extremely rare today in football although a lot more games take place\textsuperscript{164}. In fact, at the University (college) level there has been no on-field related death in the USA between 1999 and 2007\textsuperscript{165}. Although other factors than the introduction of the forward pass could also help explain the reduction in deaths in modern football (such as improved protection, better training techniques, medical science advances), the immediate impact after its introduction at least allows us to affirm that it did have an effect on the safety of the participants.

As for the second objective, making the sport more attractive to spectators, we could say that in the long term at least the rule did have the desired effect since football is one of the most popular spectator sports in the USA\textsuperscript{166}. But modern football is very different from the game that was played in the 1900’s and not only the introduction of the forward pass helped build the popularity of the sport. Like for the reduction of deaths on the field, the rule most likely played a role, but was probably not the sole and only reason for achieving the objective.

4.2.1.2. The effect of radical changes to the rules of the game on the value of players

We are therefore able to see that radical rule changes do not only modify the sport in which they are implemented, but they also have the tendency to actually create completely new sports. The changes are so important that the way the sport was played before is not relevant any more. That makes any attempt to reasonably compare the performances of a player before and after the rule change relatively difficult as the game changed in a radical way. A player may perform as good as before, better or worse after a radical rule change but this has nothing to do with a change in value due to the rule change. The whole scale has changed and the exercise of evaluating the change of the value of a player is as difficult as comparing players in different sports.

Let us look at the impact the introduction of the forward pass had on the pool of players in football. The main impact was the arrival of new players whose skills were not required in the old version of the game, but made indispensable in the new one. The obvious impact on the


\textsuperscript{165} Idem.

player pool is that you now needed players that were able to throw and catch these forward passes. In addition, teams now needed players to defend against this new offensive opportunity.

The position of the quarterback changed from a running to a passing position; it also made him the leader of the offence since he was the one handling the ball first on most plays. The receivers, the players catching the ball, are usually fast and agile but too fragile for the previous pounding style of football. These players had no interest in playing the bruising game but they were able to enter the game with the forward pass. To make that point even clearer, one could argue that football, before the introduction of the forward pass, was closer to Rugby than to the modern game we now know under its name. With that in mind, the difference in modern player physiology in both sports can make us see that the introduction of the forward pass rule opened the game to different types of players, forcing a specialisation of the players in one particular role. Modern rugby players, a game without forward passing, are less specialised than football players, smaller than the big linemen, but bigger than the small receivers. Different games require different types of players.

Unfortunately, the introduction of the forward pass is one of the very few examples of a radical rule change that one may study. Most sports have been operating with the same set of rules since the end of the XIX\textsuperscript{th} century and have not radically changed it since then.

4.2.1.3. Conclusions on radical rules changes

What can be definitively said on the impact of radical rule changes on the value of players using the limited data and example available to us from the introduction of the forward pass?

First, it is quite clear that radical rule changes have an impact on the value of players. The problem is to measure that impact. Such rule changes create a completely new scale in order to evaluate the players by creating a different sport. There are going to be many cases of players being excluded because their skills are now obsolete and now having a decreased value close to 0 as well as new players entering the player pool as their abilities, that were not

---

169 SZYMANSKI, S. AND ZIMBALIST, A., see note 57.
useful under the old set of rules, are now in demand - for these players one can say that their value increases infinitely, going from 0 to X. Radical changes will therefore trigger radical variation in the value of some players.

Second, the great difference between both sets of rules makes it difficult to effectively compare the before and after variation in the value of individual players. With a radical rule change, the points of reference and the result of the equation totally change. It therefore makes any comparison with the past very difficult. Such changes do not only change a sport, they create a totally new one that may require a completely new set of abilities, in the case of the forward pass the ability to throw the ball accurately over long distances. If the value of a player change under the new rules it is mainly due to the fact that the reference frame changes completely. One could have the same effect if the player changed sports completely. Hence, the consequences of such changes become impossible to predict with accuracy.

Finally, the comparison between the aggregate value of all players before and after a radical change is as difficult to make as for individual players. As said earlier, we are faced with a new game and any change is more a comparison between two new games exploiting different markets, so the usefulness of a comparative approach is, in this case, also limited.

Keeping in mind the goal stated at the beginning of this section, one can conclude that radical rule changes indeed affect the value of players. When the game completely changes, this creates a new scale to measure the talent of the players. The comparative approach loses some strength here since with the new set of rules we now compare two different sports that exploit two different markets. But at least we are able to say that, for the players who survive the change, radical changes in the rules change the way they are evaluated and therefore may trigger a change in their value. The usefulness of such a comparison is limited since the sports and markets exploited are radically different, but even if the value of the variation may not be useful or mean anything, the simple fact that such a variation exists allows us to see that changes in the rules of the game have an effect on the value of players.

In the next section, I will try a more measured approach examining the effect of mild rule changes on the value of players and hope that in this case a comparative approach of the state of the sport before and after the rule change will allow to learn more about the effect of the change in the rules of the game on the value of players.
4.3. Mild rule changes

In this section, I will examine the effect of mild rule changes on the value of players. The goal is to try to explain how these changes affect the value of the players. By studying these changes and taking an example from baseball as an empirical test, I will then be able to provide some general theory on the effect of a change in the rules of the game on the value of players. We saw earlier that radical rule changes resulted in some changes in the value of players but that it was almost impossible to measure the degree of change since it was so extreme. Therefore mild rule changes will hopefully help observers better understand the effect of rule changes on the value of players by providing a more precise instrument to measure it.

4.3.1. Definition of mild rule changes

In this section, I will first define what I understand by mild rule changes. These are changes that modify one aspect of the game without changing it completely or introducing a new aspect to the game. It could also be called rule adjustments or rule clarifications. The situation after the change is still recognisable and the sport is still played in more or less the same way.

4.3.2. Justification for mild rule changes

As said earlier, rule changes are relatively rare in sports. Rule stability may be one way of helping competitive balance as all participants know the rules since they began playing the sport and know that they will remain the same for a while. When any change happens, the justifications put forward by the authorities of the sport are usually the same: to improve the safety of the participants and the interest of the spectators. This section should allow determining if, in addition to these stated goals, these changes have an impact on the value of players. Logic dictates it that both stated effects should be positively correlated or at least linked to the value of players. If the rule changes do result in fewer injuries and more spectators, then the value of the players should be positively influenced, too: they will have a longer playing career due to fewer injuries and be able to ask for higher salaries with more

---

spectators giving more revenues to the sport and making more money available for the players. If the rule changes have different effects on the value of the players, sometimes positive, other times negative, then they are other effects that are not disclosed or known by the sporting authorities. But let us first look at the stated goals of the sporting authorities when they adopt mild rule changes.

4.3.2.1. Safety of the participants

Safety of the participants was discussed at length in section 4.2.1. regarding radical rule changes. For mild rule changes the goal is the same, but the means taken are less radical and do not create a completely new game. A recent example of such a change includes the automatic expulsion of a player who commits a foul on a sliding tackle from behind in soccer. Although the main effect of such rule changes may be to reduce the injury rate of a sport, it probably also helps when a secondary effect is to polish the image of the sport, making it more commercially attractive. In fact, the “safety of the participants”-justification sometimes seems to be just a sub-category of the commercial success of the game. The idea is that by improving its image, the sport will become more attractive to spectators. The fact that players may be able to have a longer playing career because of the lower rate of injury is almost just an afterthought. Let us examine more closely the second justification that seems to be the more important one, the interest of spectators.

4.3.2.2. The interest of spectators

The main motivation behind the changes in rules seems to be to make the game commercially more appealing by boosting either stadium attendance and/or television ratings. In economic terms, one can say that the objective of the rule change is to increase the demand for the sport. The hope is that with new rules, more people are going to watch the sport, bringing more money in the league’s coffers. The impact of this on the value of players is relatively straightforward: by having more demand for the sport, the aggregate value of players should be higher as teams have more money to pay them. This requires the simple assumption that the percentage of the revenues devoted by teams to players’ salaries remains the same. The same share of a bigger pie will result in more money for the players. But does that hold in reality?

Following the discussion regarding the sportsman owner in the previous chapter, this assumption seems rather solid. The difficulty of the exercise here is to make sure that the change in rules is the cause of the higher revenues.

4.3.2.3. Lowering the mound to get higher scores in Major League Baseball

Let us now consider one example that will allow us to see more clearly the real impact on a mild rule change on the value of players. I chose that example since I think that the availability of individual players’ performance statistics in baseball will allow me to better detect the effect of the rule change on their on-field performance. Also, rule changes that result in detectable individual on-field performance changes for the players are relatively rare as the different sports try to keep their on-field rules as stable as possible. I am confident that the study of the 1968 MLB rule change will allow making general conclusions on the effect of mild rule changes on the value of players that are going to be useful for all sports.

There are few more blatant attempts in changing the rules to get more demand for a sport than what happened in MLB in 1968. That year, the league decided to lower the elevation of the pitcher’s mound from 15 inches (around 40 cm) to 10 inches (around 25 cm). The objective was to help the offensive side of the game by making the ball easier to hit and therefore producing more points. The reason why such a change in the elevation of the mound should help the batter is that it reduces the pitcher’s effectiveness. This can be explained by a physic principle called the “gear effect”. It states that when throwing an object with an arc, the wider the arc, the greater the velocity of the object for the same force. Throwing off a lower mound will result in a smaller arc and a lower ball velocity for every pitch.

Assuming that slower balls are easier to hit, slower pitches should make the job of the batter easier and result in more points scored. The assumption of the MLB authorities was that this added offence was deemed to stimulate demand for the game since it is assumed that fans want to see points scored. According to that thinking, points generate excitement and

---

172 See section 3.3.
enhance the enjoyment of the game for fans that are then willing to give more money to follow their favourite sport and this will result in higher revenues for the teams. Again in economic terms, the lowering of the mound should result in a greater demand for baseball games. But did it really result in more points scored?

4.3.2.3.1. Effect of the lowering of the mound on individual baseball offensive statistics

The lowering of the pitchers’ mound provides us with a rare opportunity to objectively see the effects of a mild change of the rules of the game on the on-field performances of sportsmen. This is due to the nature of the sport of baseball as a team sport but having more in common with an individual sport with its most important phase, the pitching and hitting part, being up to two players facing each other. In that phase we have one player, the pitcher, who stands alone and measures himself with another individual, the batter. As a result of the mild change in the rules of the game, these individual performances can be measured in order to compare the effects of the change on the individual players. There is, in general, no influence of other players that could falsify the results during that phase.

When one compares the statistics for the 1968 and 1969 seasons, the effects of the rule change seemed instantaneous since there are no doubts that more points were scored in 1969, after the rule change. The 18 teams that took part in both seasons scored in average 0.54 more points per game. Using individual player statistics, each player participating in both seasons scored in average 1.5 more runs in 1969. These numbers seem to confirm the fact that the lowering of the mound did indeed stimulate the offensive side of baseball.

4.3.2.3.2. Is the lowering of the mound the only explanation?

But only considering the lowering of the mound does not give us the complete picture of the changes that occurred between the seasons of 1968 and 1969. In 1969, MLB also added four new teams to their leagues (two in the National and two in the American). These teams allowed more players to have access to MLB, players who had not been good enough to play in the league the year before. So there was a dilution of the pool of talent, especially for pitchers, that may also partly explain why more points were scored in 1969. To explain this

---

177 Idem.
further, the lowering of the pool of talent in pitchers is the most important effect of expansion. Pitching has a disproportionate importance in the sport and in the production of points since the pitcher is implicated in approximately three quarters of the action in a game. A batter for his part will have around three to five appearances in a game. This seems to reflect the old saying in baseball “The name of the game is pitching”.

Even assuming that the rule change was responsible for the additional points scored in 1969, what about the long term effects of the change? How can they be measured?

An extensive study on the impact of expansion in MLB\textsuperscript{178} has shown that the average amount of points scored in MLB per team per game was the same in 2004 as in 1950. It was found that in the first years following the addition of a new team, more points per team were scored, but in the long term, the amount of points scored in baseball remained constant in the last half of the century despite all expansions and rule changes. An interesting side note to that study is that the authors have chosen to “control” for the uniqueness of the 1969 season by excluding it from their study, because of the size of the expansion (16% more teams) and the lowering of the pitcher’s mound.

These results are very useful for the purpose of this work even if the studied aspect is different. It shows that over almost half a century, the average amount of points scored in baseball remained constant. This means for this thesis that the lowering of mound did not have, over a long period, an influence on the amount of points scored. It also confirms that any change in the distribution of talent in baseball, due to expansion or to a change in the rules, will result, in the short term, in a change in the number of points scored.

Considering that over a long period the impact of a mild rule change seems to disappear or at least is difficult to detect, it forces us to conclude that the motivation of the sporting authorities adopting the change may have little to do with the real impact the change will have in the long term.

\textsuperscript{178} QUINN, K. AND BURSIK, P., see note 105.
The next question that needs to be answered is how it is possible that over half a century in which the sport evolved in so many different ways, the average amount of points scored remained the same.

The impact of a rule change that benefits the offensive side of baseball seems to be the same as concerning the expansion: it results in a temporary boost in points scored but that effect dissipates on the long run. How can that be explained? In the next section, I will try to answer that question by examining closely how players react to rule changes and how it affects their value.

4.4. Players and changes to the rules of the game

Changing the rules of the game even in a very trivial way can have unforeseen effects on the players. They have invested years of training in skills to use with the old set of rules and now they have, in a short time, to adapt to new ones. Even if the rule changes are mild, they may affect the players’ skills in an important way. In this section, I will first examine the effects of a rule change on the on-field performance of a player. After that, I will look at its consequences on the individual value of a player and, finally, the effect of the rule change on the aggregate value of all players in a sport.

4.4.1. Effect of a change in the rules of the game on the on-field performance of players

A change in the rules of a sport will have three possible consequences on the on-field performance of a player: 1- the new rule will have no effect on his performances, 2- The new rule will provide the player with an advantage and thus make him a better player and help improve his performances, 3- The new rule will handicap the player, making him a worse player and resulting in a decline of his on-field performances.

These are all ex post evaluations of the effect of a rule change since at the very moment the rule is changed, it is impossible for the player to know ex ante with certainty what is going to be the influence of that change on his performances. He may guess the effects but there are no certainties until he starts performing under the new rules. The 1968 rule change in Baseball was supposed to disadvantage pitchers, so a pitcher could probably predict that his job would become tougher at the start of the 1969 season. However, he did not know how tough it would
be and how the other pitchers would perform under the new set of rule. Maybe his abilities predisposed him to be as good as or even better than before, but he had no means to know it before pitching under the new conditions.

Additionally, the player has little or no control on how the rules are changed since the decision is usually taken unilaterally by the league or sporting authorities. Even if the players are lucky enough to have an influence on the league’s decision making process, it rarely goes beyond a consultation with some union representatives. So it is therefore more appropriate to focus on the *ex post* aspects of the change of the rules. Also, the traditional way of determining the value in sports has always been to consider the past performance of the player to determine his value, so using the *ex post* reasoning in studying the rule changes helps to keep some kind of constancy.

### 4.4.2. Effect of a change in rules of the game on the value of players

Keeping in mind the above mentioned three possible effects of rule change on a player, it is now time to study the effects of a rule change on the value of players. In order to make some general conclusions that will be applicable to all sports, we will have to look at the effect of a rule change but under the three following assumptions. These assumptions are relatively robust and valid when we look at the effect of a rule change on the value of players.

First, I will assume that the amount of money available to pay the players’ salaries stays constant before and after the change. This assumption means that either the revenues of the league stay constant or that the percentage of revenues attributed by the teams to players’ salaries stays constant. The goal here is to insulate the value of players against a rise in revenues of teams in order to clearly see the net impact on the players.

The second assumption is that the individual performances of the players are relatively easy to measure. It may be an assumption that is more robust in some sports like baseball than in others. But without it, there would be no way to see the impact of the rule change.

The final assumption is that the remuneration of a player is positively correlated with his performance on the field. The better he performs, the higher his salary will be. This

---

179 LEWIS, M., see note 149.
assumption is in accordance with the empirical findings made in the sport field\textsuperscript{180} and is also rather robust.

In order for a change in rules to have any influence at all on the value of the players, the players’ salaries have to be linked to their performances on the field. Only when such a link exists do we have a redistribution of wealth between the players.

I think that these assumptions are relatively easy to accept and reflect, up to a certain degree, the situation that exists in most professional sports.

\textbf{4.4.3. Redistribution of players’ value}

Building on these assumptions and empirical data that I exposed in the preceding section, we can safely make the following conclusions as for the short term of rule change.

If the revenues of the league remain constant or at least the percentage of revenues allocated to players’ salaries stays constant following the rule change, then the change in the value of players will end up being a purely redistributive exercise. Some players will end up getting more money at the expense of others since they will perform better under the new rules. For some players, the change in the rules will not affect their financial situation since their performances will not change. The whole exercise will shift money from one group of players to another. The relative value of some players may change due to the redistribution, but the aggregate value of the player pool will stay the same. If the change in the rules of the game does not result in an increase in revenues or of the share of these revenues that is paid to the players, they will end up receiving the same share of revenue but differently distributed. Individual players may see their value change, mainly due to pure luck, since nobody can know \textit{ex ante} how he will adapt to the new rules. Pure luck could also be involved in an aggregate value increase since there does not seem to be a clear relationship between the change in the rules and the increase in revenues of a sport (as we have seen earlier, the demand for sports is also influenced by other factors\textsuperscript{181}).


\textsuperscript{181} As seen in Chapter 2.
But the effect of a change of the rules of the game on the redistribution of value of players is likely to be only temporary. On the long term, the value of players should not be affected by the rule change since such a change does not constitute a change anymore when players have been playing their whole career with these “new” rules. The impact of the rule change will dissipate over the years. The rational reaction of the players having started their career with the new rules will be to optimise their skills and abilities in accordance to these new rules. For them, the rule change is not a change at all since they have been playing with it since the start of their career. Over a long period, these players will replace the actual ones who had to live through the rule change as they retire or are fired because their skills have become obsolete. When the current player pool has been completely replaced by new players, the mild rule change will then cease to have an effect on the value of individual players since it has been crystallised in the training and economic system of the game. The impact of a change to the rules will then have dissipated as the experience of MLB has shown in the previous section.

To conclude this section we can say that a change in the rules of the game has a temporary redistributive effect between the various players participating in the sport at the moment this change is enacted. Assuming that all other aspects of the sport remain constant, when the generation of players contemporary with the rules changes has been completely replaced by other players who have played with the new rules all their careers, then this redistributive effect disappears, having been crystallised by the development system of the sport and taken into account in the valuation of these players. At this point, it does not represent a change any more.

However, there are cases in which rule changes lead to value changes that are not merely redistributive ones. I will try to look at them in the next section.

4.4.4. Change in the aggregate value of players

The only way a change in the rules of the game will influence the value of the players is if the revenues of the league do not stay constant after the rule change and if the new rules generate a greater demand for its product. With greater revenues for the teams the players will get higher salaries, unless we assume that the teams will reduce the proportion of their revenues that they assign to players’ salaries. This is however unlikely since the salaries of players have historically grown at a rate at least equal to, often even greater than the revenues of their
sports\textsuperscript{182}. In addition, as shown by the sportsman owner model, teams are not pure profit maximising entities and have the incentives not to reduce the proportion of the revenues that they assign to salaries\textsuperscript{183}. So the result of higher revenues will most likely be a greater aggregate value for the players.

We can safely assume that if the rule change results in an increase in revenues for the sport in general, the percentage of revenues allocated to the players will stay the same or be even greater than before the rule change. Assuming that this growth is relatively equally distributed to all teams, the consequence of growing revenues will therefore be a growing aggregate value of players. The way in which it is going to be distributed between the players will be similar to the effect of a rule change without growth in revenues: a redistributive effect of that growth in value according to the influence of the rule on the on-field performance of the player. This redistributive effect will be identical to the one that has been discussed earlier in this chapter. In the case of a growing aggregate value of players, there will be an internal redistribution of the individual shares that each player gets from the new aggregated value. This will be done, like in the previous case, in accordance to how the new rules affect the on-field performance of the player. So even in the case of growing revenues, the redistributive effect will be present.

For the players entering the league after the influence of the change has been crystallized, unless the revenues of the sport grow and affect the aggregate value of players, they will not feel any impact from the change. The benefit or penalty of the rule change will already be included in their value.

\textbf{4.5. Conclusion}

What is, in conclusion, the impact of a change in the rules of the game on the value of players? As we have seen in this chapter, we can divide that impact in three categories.

The first one covers the impact of radical changes in the rules of the game. In that case, we do not have a simple change in value, but we have a change of the complete scale upon which the


\textsuperscript{183} See section 3.3.
players’ value is measured. With radical changes, the variation in value is unpredictable and impossible for the players to anticipate, for them this is the equivalent of having to play a different sport. Therefore, for the purpose of this work, we can say that a radical change in the rules of the game results in a change of method of the valuation of the players and therefore makes any comparison before and after the change very difficult. One must be careful and remember that this is a comparison closer to a comparison between two different sports than to one between two sets of rules for the same sport.

The second category covers the short term effects of a mild change in the rules of the game. In this case, when looking at the individual value of players, we can see a redistribution effect between the different types of players. The players benefitting from the changed rule will get the value that is lost by the ones that are disadvantaged by it. If the change also results in a growth in revenues of the sport, then, under certain conditions, this will result in an increase in the aggregate value of the players.

The third category covers the long term effect of the mild changes in the rules of the game. For the purpose of this category, long term means a change in generation: when all the players that were active when the change was implemented are replaced by new ones that have been playing with the new rules for all their careers, this is what constitutes long term. In this case, the change does not constitute a change any more; it has become part of the standard rules. Therefore, any influence that could have been exerted on the value of this new generation of players has been integrated in the new formula that is used to determine their value.

Now that we were able to determine the impact of a change in the rules of the game, the next step will be to examine what sort of impact a change in the rules of the commercial game of a sport has on the value of players. This is the exercise that I will carry out in the next chapter.
Chapter 5 - The effect of a change in the rules of the commercial game on the value of players

“Gentlemen, we have the only legal monopoly in the country, and we're f--- it up.”
- Attributed to MLB Atlanta Braves owner Ted Turner

5.1. The rules of the commercial game

Although generally not well known and understood, the rules of the commercial games have the most important influence on the value of the players. They could also be called the internal rules of sport since they are mostly regulations adopted by the sport authorities themselves in order to regulate some specific commercial aspect of a sport. They generally have been considered fair game for judicial revision, like any other internal rule of a private entity. The rules that I will examine in this thesis are the ones implemented in order to regulate the relationship between the players and the teams. The remaining rules, the ones regulating the relationship between the teams themselves, will not be a subject of this thesis since they do not exert such a direct influence on the value of the players as the ones regulating the relationship between the players and the teams. I however tried to account for these rules in chapter 3 when discussing the way that teams are valued.

The main influence that the rules of the commercial game have on the value of the player is when they restrict the players’ freedom of movement. By doing so, the rules directly influence the market for players and in consequence their value. There are other rules of the commercial game that also influence the value of players but looking at the restrictions to the movement of players will allow to drawing clear conclusions that will make it possible to make general statements about the other rules of the commercial game. By looking at the rules that directly affect the value of players, it will be possible to deduce the effect of the ones that affect it indirectly.
Therefore, I will examine how the restrictions on the movement of players have evolved in two sports, American football and soccer, and how they affect the value of the players in these two sports.

5.1.1. Regulations on the free movement of players

The oldest and most controversial regulation to affect the free movement of players is the reserve clause. The way it is implemented may vary depending on the sport but the result is similar, it limits the players’ ability to look for the team that values him the most by forcing him to stay at the service of one single team, the one that first recruited him.

5.1.1.1. The reserve clause

The reserve clause was “born” at the start of the 1880 baseball season. The National League unilaterally added a clause to its standard player contract that allowed a team to unilaterally reserve the services of a player for the following seasons at the expiration of his contract. Since all contracts had a one year term, it meant that the player was forced to remain with that team indefinitely. At first, this was seen as a status symbol by the players because the number of players that could be reserved was limited to five per team. But, at the start of the 1883 season, the league expended it to all players. The players rightly saw this clause as a means to hold their salary down by denying them the right to offer their services to all teams, but the owners justified it by claiming that it was adopted in order to preserve competitive balance since it helped poorer teams retain their good players. Without the reverse clause, they argued, all the good players would end up playing for the richer teams and this would result in an unbalanced league where only rich teams would succeed. But what was the real impact of the reserve clause? Did it actually influence the competitive balance of baseball?

The effect of the reserve clause on competitive balance is hard to assess for even in a free market for the services of players, there will still be some restrictions since as stated in chapter 1 sport organisations usually operate as cartels. As a result, the nature of the market for the services of the players will be limited to the members of the cartel. Even today the reserve clause has survived in a negotiated form in most sport organisations. It now takes the

184 SZYMANSKI, S. AND ZIMBALIST, A., see note 57.
185 AHLBURG, D. AND DWORFIN, J., see note 182.
186 Chapter 1, section 1.3.3.
form of other restrictions of the freedom of movement of players. Teams continue to justify these restrictions in the same way they did in 1880, saying that such clauses are necessary in order to protect the competitive balance of the league. But it became clear that this argument does not stand serious examination. It has been known for over 50 years\(^{187}\) that the reserve clause does not prevent the best players from migrating to the richest teams. Players always ended up in the teams that valued them the most irrespective of the limitations to their movements. So it is false to say that the reserve clause helps the competitive balance of leagues. Instead, the effect of the reserve clause is to allow the teams to capture most of the monopolistic rent coming from the operations of a sport league by paying their players under their market value.\(^{188}\) This is done by restricting the free movement of players by preventing them from offering their services to all the teams in the league. This reduces the players’ negotiating power since they are forced to negotiate only with one team. It also limits the competition between teams for these talented players. With the reserve clause, teams cannot try to acquire the players by offering them better conditions; their only choice is to negotiate with the team that has the rights to the player. To use the formula introduced in chapter 3\(^{189}\), the reserve clause raises the AC of the player. Since the willingness to pay of the teams stays the same, the reserve clause results in the teams taking part of the value of the players’ talent away from them. In addition, the reserve clause puts a downward pressure on the salary the player is able to negotiate with his current team since his only option in the case of an unsatisfying offer is not to play at all. He has no right to try to get other competitive offers from other teams. This forces him to accept whatever is offered to him, putting him at the mercy and good will of his team.

It is therefore not surprising that the first thing the players did was to fight against the reserve clause to get some kind of freedom to negotiate with all teams.

5.1.1.2. Free agency

The reserve clause could be seen as the root of most labour problems of professional sports. Most legal confrontations between players and teams can be traced back to it. The players and their representatives seemed to have understood that this clause was one of the greatest

---


\(^{188}\) Idem.

\(^{189}\) Chapter 3, section 3.5.3.
constraints on their revenues and fought hard to eliminate it or restrain its application as much as possible. Two of the most representative labour conflicts aimed at limiting the application of a reserve clause on players were the Bosman\textsuperscript{190} case in soccer and the long judicial crusade of the National Football League Players’ Association (NFLPA) in football. Analysing these two conflicts will allow us to better understand the effect of a change to the commercial rules of the game on the value of players. The reserve clause represents the perfect example of a commercial rule that regulates the relationship between the players and the teams on a purely commercial aspect. By examining how a change in the application of this rule will affect the value of players it may be possible to come up with general conclusions regarding the impact of the commercial rules of the game on the value of players.

Both cases that I am going to examine here started with a unilateral imposition of the reserve clause and ended with the player gaining a greater degree of freedom of movement. However, the form it has taken is totally different in both cases for the NFLPA freedom was granted in the creation of a business partnership with the teams and was a negotiated solution whereas for the soccer players, a new system was imposed on them anew, without negotiation. But both systems influenced differently, the value of the players involved in the sports. Let us first look at the long judicial saga of the NFLPA, before putting it side by side with the situation in European soccer.

\textbf{5.2. The players vs. the NFL}

Even if the players of the NFL were unionized early compared with other professional sports\textsuperscript{191}, they never managed to find a negotiated solution to end the application of the reserve clause. It took the intervention of the tribunals to help them gain their freedom of movement. The judicial history of the liberation of the professional football players can be summed up in the three following decisions.

\textbf{5.2.1. Radovich vs. NFL}\textsuperscript{192}, the applicability of antitrust law to football

Mr. Radovich was a player for the NFL Detroit team and asked, for personal reasons, to be transferred to the Los Angeles team but his request was refused. As a result, he unilaterally

\textsuperscript{190} Bosman, see note 9.
\textsuperscript{192} Radovich, see note 9.
terminated his player contract and went to play for the Los Angeles team of a rival football league. When he tried to come back to the NFL after the failure of the rival league, no team offered him a contract although he was considered a very talented player. He had been in fact black-listed by the NFL authorities and teams feared severe sanctions from the league if they hired him. Radovich sued the NFL using US antitrust laws, arguing that it entered into a conspiracy to monopolize and control organized professional football. He added that this conspiracy resulted in a boycott of his services, enforced by the fear of teams to suffer sanctions from the league if they disobeyed the authorities of the league. By boycotting him they prevented him from earning a living as a football player.

The main question that interested the court in this case was to determine if the business of professional football was covered by the antitrust laws. It was an issue at that time since the US Supreme Court had, in 1922, declared that professional baseball was exempted from the antitrust laws\(^{193}\) by ruling that baseball did not constitute interstate commerce. According to US anti-trust law, an industry can be subject to the law only if it is commerce between different states. The NFL tried to get the same exemption and asked for an early dismissal of Radovich’s complaint. The court found that professional football was in fact covered by the antitrust laws since it constituted interstate commerce and consequently that Radovich was allowed to use them to contest the league’s rules. Although the decision does not tell us what happened to Radovich’s complaint, it is a very important decision as it showed that football was subject to the antitrust rules and provided the instruments that have been used by the players in subsequent actions.

### 5.2.2. After Radovich

After the decision, the league decided to grant some liberty to the players. But they still remained greatly restricted by what was called the “Rozelle rule”\(^{194}\). This rule established a certain procedure that a player had to follow in order to escape the application of the reserve clause. It took the form of a two steps procedure: first, the player who wanted to get his freedom was required to play the last year of his contract at a reduced salary and second, when he signed his new contract with a new team, his new team was required to pay a compensation to his former team for the loss of the player. If the two teams were unable to

---

\(^{193}\) *Federal Baseball Club of Baltimore, see note 2.*

\(^{194}\) In "honour" of the NFL Commissioner at the time, Mr. Pete Rozelle
agree on the compensation due, the commissioner would determine it. Compensation rarely took the form of money; it took the form of draft picks.\textsuperscript{195} The penalty for signing a player could endanger the future chance of the team to compete if the new player did not meet the team’s expectations. The commissioner’s decision was final, without appeal and discretionary since there were no guidelines to objectively determine the compensation. This system led to a market in which few players were willing to make the monetary sacrifice to gain their liberty and even fewer teams were ready to sign players without knowing in advance what would be their price.\textsuperscript{196} The players, in order to gain more freedom, decided to contest this rule, too, using antitrust law again.

5.2.3. \textit{Mackey, the fall of the Rozelle rule}

In the case \textit{Mackey vs. NFL}\textsuperscript{197}, the players contested the Rozelle rule by arguing that it was an abuse of a dominant position, prohibited by the US antitrust laws. The NFL argued that the rule was protected by the non-statutory labour exemption that allows parties to act in a way prohibited by antitrust laws if it is done in an arm’s length labour negotiation.\textsuperscript{198} This exemption acknowledges the fact that some anti-competitive behaviour done in a labour relation context can maximize the social welfare and should be allowed. In this case, the trial judge and the appeal Court did not agree that the Rozelle rule was the result of arm’s length labour negotiation and ruled in favour of the players declaring the Rozelle rule illegal. None of the courts actually tried to determine if the rule resulted in an increase of social welfare, but they stated that limitations to competition in professional sports were not illegal \textit{per se} and that a court should apply the rule of reason when faced with such cases.

5.2.4. After \textit{Mackey}

Before the US Supreme Court could rule in the \textit{Mackey} case, a negotiated settlement was reached between the parties. The settlement still allowed the teams to have a right of first refusal and compensation, based on a tariff written in the CBA, when a player changed team. The Rozelle rule was abrogated but was replaced by a new rule that had the same effect and

\textsuperscript{195} See section 5.4. for a more details on the NFL draft.
\textsuperscript{196} Data on the difficulty of movement of players was presented in the Mackey case, see 5.2.3.
\textsuperscript{197} 407 F.Supp. 1000 (D.Minn 1975), appeal: 543 F.2d 606 (8th Cir. 1976).
\textsuperscript{198} For more details on the implementation of that exception in the professional football context, see \textit{Brown v. Pro Football Inc.}, 518 U.S. 231 (1996) and for a general explanation of the exemption see \textit{Connell Const. Co. v. Plumbers & Steamfitters Loc. U. No. 100}, 421 U.S. 616 (1975).
that was even better protected from antitrust law since it was covered by the antitrust non-statutory labour exemption\textsuperscript{199} as the players found out when they tried unsuccessfully to challenge the new rule. So even if the players had won a victory in court, they would have won, in fact, little freedom because the union made some bad judgments during the negotiations. After the expiration of the CBA the players tried to get rid of that rule but the teams refused and answered by taking unilateral action that restricted even more the freedom of movement of the players. These measures, called Plan B, ended up being also contested in court by the players, again using the antitrust rules. In order to avoid the non-statutory labour exemption, the players’ union disaccredited itself and a group of individual players sued the NFL.

5.2.5. McNeil vs. NFL\textsuperscript{200}, the final judicial act

Mr. McNeil and eight others players contested Plan B as a group non-affiliated with the disbanded players’ union. Their main argument was that the rule prevented them from freely negotiating their contracts with the highest bidder and therefore constituted an abuse of a dominant position from the league. The teams’ owners claimed on their part that everybody gained from that system and that financial collapse was imminent if the rules were changed because it would allow salaries to escalate and destroy all the profits of teams. In addition, they argued that it was crucial to maintain the competitive balance of the game since without this rule only rich teams would be able to hire the best players. In the end, the court ruled in favour of the players, deciding that the Plan B rule was an illegal restriction to the market for the service of players and that it had a substantial harmful effect on their compensation. However, it recognized that the system helped to maintain the competitive balance of the league but that it was more restrictive than necessary, the same goal being reachable using a less drastic form of regulation\textsuperscript{201}. But this last conclusion is questionable since, as shown earlier, it has been shown that reserve clauses have no effect on the competitive balance of a sport league\textsuperscript{202}.

This case was a great victory for the players who ended up being able to offer their services to the highest bidder. Although the McNeil judgment could have allowed a completely free

\textsuperscript{202}EL-HODIRI, M. AND QUIRK, J., see note 187.
market for players, again the players allowed restrictions to their rights to be included in the CBA signed in 1993. Still, the decision allowed them to use their newly acquired freedom as a bargaining chip in exchange for a greater share of the revenues of the league. Both parties seemed to have realized that a partnership had the potential of being more profitable than these repeated judicial jousts.

5.3. The negotiated settlement

In the following section, I will examine the sections of the NFL CBA that restrict the free movement of players as well as the compensation scheme which is also a corollary of the reserve clause and seems to have been adopted in exchange for the restrictions. This is the negotiated result of the judicial saga that opposed players and teams. If one wants to understand the impact of such restrictions on the value of football players, a careful examination of these rules is necessary.

5.3.1. The salary cap

The first component of this settlement is a restriction on the total amount of money a team can spend on the salaries of the players it employs. This is called a salary cap. Under the CBA, the teams in the NFL operate under the constraint of a salary cap. This is a “hard” cap, which takes into consideration all the advantages given to the players to make sure that teams do not use non-monetary compensation to get around the restrictions. The amount of the cap is established at the end of each season and corresponds to a certain percentage of the league’s predicted revenues for the next one. The CBA provides an extensive definition of revenues in Article XXIV, section 1 that includes and excludes specific types of revenues but for the sake of simplicity, it is possible to say that the cap gives the players the right to around 50% of the total revenues of the league. But the players had to pay a certain price to achieve that.

204 For the 2008/2009 season it was 57.5%. Source NFL Players Association website: http://www.nflpa.org/Members/main.asp?subPage=CBA+Extension+Features, date of download: February 2009.
5.3.2. Contractual uncertainty

A very important aspect of the CBA is that the players’ contracts are not guaranteed. This means that a player can be fired at any time during the course of his contract; a three years contract may end up having an uncertain duration for both parties. On the positive side, once fired the player is free to go to any other team to offer his services, irrespective of the limitations imposed on the free movement of players by the CBA. However, this may not result in the optimal situation for the player. Such an event may happen at any moment during the season and may happen when the demand for the service of the player is low or non-existent like at the end of the season. This is hardly a way to help a player maximize his revenues and more a way of treating an unemployed worker.

5.3.3. The negotiated reserve system

As previously stated, the CBA restricts the mobility of players. The extent of that restriction depends on the CBA category the player falls in. There are four such categories\textsuperscript{205}:

1- Rookie non-free agents
2- Restricted free agents
3- Unrestricted free agents
4- Special cases (Franchise and Transition players)

5.3.3.1 Rookie non-free agents

There is no definition of rookie non-free agent in the CBA. It can be established \textit{a contrario} by saying that any player not fulfilling the criteria of any other category. In fact that means that a player with less than three seasons of experience is a rookie non-free agent. The rookie non-free agents have no freedom of movement at all. They must sign with the team that drafted them\textsuperscript{206}. This total absence of freedom can be rationalized by the fact that without this, the college draft would be meaningless since teams would lose the rights on a player at the end of his first contract. That would mean that any rational player would only sign a one year contract with the team that drafted him in order to be able to offer his services to all teams the

\textsuperscript{205} Article XIX of the NFL CBA.
\textsuperscript{206} More on the draft later in section 5.4.
year after and be freed of the obligation to play for the team that selected him. The system therefore protects the rights of teams on rookie players and could also be seen as an incentive to train and educate young players. If they were able to leave at the end of their first contracts, there would be no incentive for teams to invest any money to make sure that the players reach their highest potential. The rookie non-free agents are subject to a complete restriction on their movement; as said above, they can only play for the team that drafted them.

5.3.3.2. Restricted free agents

They are defined in the CBA as “any Veteran player with three or more Accrued Seasons, but less than four Accrued Seasons.” A player that meets these terms and has no valid contract at the end of a season is free to negotiate a contract with any team in the league, subject to certain restrictions.

At the end of the season a team should make an offer to all its restricted free agents and the amount of that offer will place them in one of the following four categories.

A) If a one year 250,000 $ salary contract is offered, then the team gets a right of first refusal on the player.
B) If a one year 275,000 $ salary or 110 % of the player’s previous year salary, whichever is greater, contract is offered, then the team gets a right of first refusal on the player and a draft selection at the player’s original draft round from the new team as compensation if this right is not exercised.
C) If a one year 600,000 $ salary or 110 % of the player’s previous year salary, whichever is greater, contract is offered, then the team gets a right of first refusal on the player and one first round draft selection from the new team as compensation if this right is not exercised.
D) If a one year 800,000 $ salary or 110 % of the player’s previous year salary, whichever is greater, contract is offered, then the team gets a right of first refusal on the player and one first round draft selection and one third round draft selection from the new team as compensation if this right is not exercised.

---

207 Paragraph 2(a) of Article XIX of the NFL CBA
208 Paragraph 2(b) of Article XIX of the NFL CBA. The reader should note that the amounts used in that paper are the ones originally written in 1993 in the NFL CBA. Paragraph 2(e) of Article XIX allows an increase of these amounts to account for inflation, but I will ignore it for the purpose of simplicity.
In all except one of these categories there are two restrictions to the mobility of the players: 1) the right of first refusal and 2) draft picks compensation, coming from the team signing the free agent. As a trade-off for these restrictions the player is guaranteed a minimum salary that is linked to the importance of the restrictions to his free movement. The greater the restrictions, the greater his salary will be. I will now examine these restrictions in detail.

5.3.3.2.1. The right of first refusal

The right of first refusal allows the former team to match any offer accepted by the player from another team in order to keep him. This limits the mobility of the player who may have to stay with his old team.\(^{209}\) However, the player is compensated by having the same financial conditions that he would get if he was allowed to leave the team.\(^{210}\) This seems to be a winning situation for both parties. The teams are allowed to keep their players and the players are able to get higher salaries. The empirical data confirms that: between 1993 and 2001, only 5% of the restricted free agents changed teams, but they got an average increase of their yearly salary of 90%.\(^{211}\)

5.3.3.2.2. Compensation

If the old team does not exercise their right of first refusal, it is awarded draft picks from the new team in the next College draft. This is the remnant of the Rozelle rule and has the same effect of restraining the movement of players. But the players freely negotiated it and have received in return a guarantee of a higher salary for the affected players.\(^{212}\)

5.3.3.2.3. The effects of restricted free agency on the value of players

Once the player has completed his third season, he is able to use the market forces to negotiate a higher salary. He is not totally free to change teams but the restrictions on his movement are compensated by the possibility of getting a substantial increase in salary. The team can force the player to remain with it, but has to pay the market value of the player in

\(^{209}\) Section 3 of Article XIX of the NFL CBA.
\(^{210}\) Paragraph 3(e) of Article XIX of the NFL CBA.
\(^{212}\) DUBERSTEIN, M.J. see note 211 and Paragraph 2(b) Article XIX of the NFL CBA.
order to do so. However, the remuneration offered to a restricted free agent will always be under his real market value since any new team will have to take into consideration the draft pick(s) it has to give up in order to secure the services of the player. This will reduce their willingness to pay for the player. This reduction must however be compensated by the old team. The higher the non-monetary compensation it wants to claim, the higher the initial salary offer made to the player should be. The team must, when faced with a restricted free agent, try to guess the market value of the player in order to make an initial offer that is going to maximize their chances of retaining the player at the lowest possible cost. They have to make an offer to the player that is going to be under his market value, but combined with the restrictions of the CBA, this offer will be over the willingness to pay for the services of the player of all the other teams in the league.

For the player, the presence of a non-monetary compensation means that his salary will always be lower than the one he would be able to get on a free market. The team that wants to hire him will also have to include in its willingness to pay the non-monetary compensation that it has to give to his old team. What remains will be given to the player in salary. But the player, having access to an imperfect market, is nevertheless better off than not having access to a market at all. As the data from the NFLPA confirms, he will be able to get a substantial increase of his salary.

The consequences of these restrictions are therefore clear. First, it allows the teams to retain the services of the restricted free agents and second, it allows teams to capture a part of the value of the player for themselves through the non-monetary compensation that they get when a player leaves. This takes the form of a lower salary for the player than the one he would be able to get on a free market. In conclusion, even if it was possible to argue that these restrictions help the competitive balance of the league by allowing teams to retain their players or getting compensated if they leave, the concrete impact of these restrictions is a limitation of the salary of the players.
5.3.3.3. Unrestricted free agents

In this category we find “any player with four or more Accrued Seasons”\textsuperscript{213}. It also includes players that were fired by their teams and undrafted rookies. These players are free to contract with any team in the league. So in their cases the market really determines the value of the salary of the player since there are no restrictions of movement and they are allowed to sign with the highest bidder. Empirical evidence shows that the rules of the market really apply to the unrestricted free agents: from 1993 to 2001, 52% of unrestricted free agents changed teams and their average yearly salary increase, independent of the fact that they changed teams or not, was of about 72%\textsuperscript{214}.

5.3.3.4. Special cases: the franchise and transition player designation\textsuperscript{215}

There is however an additional restriction that can be used to limit the movements of some unrestricted free agents. According to the CBA, each team has the right to designate one unrestricted free agent as Franchise Player and another one as Transition Player. The effect of this designation is that a Franchise Player is subject to the right of first refusal and a non-monetary compensation of two first round draft picks, if the right of first refusal is not exercised. The Transition Player designation only allows the team to have a right of first refusal on the player.

To compensate for these restrictions, the designated players put in these categories are awarded a one year contract with a yearly salary that should be either the average of the five, in the case of Franchise Players, or of the ten, in case of the Transition Player, highest paid players at his position\textsuperscript{216} or 120% of his previous year salary, whichever is the highest. Again, the restriction on the player’s mobility is compensated by higher salaries. This clause is intended to be used only for the best players in the league since using it will make the player one of the best paid in his position. This should in consequence give him a salary close to his market value. But this remains an imperfect instrument since the salary will not be based on his own value but on the ones of other players. Finally, there is no guarantee that the salary

\textsuperscript{213} Paragraph 1(a) of Article XIX of the CBA.
\textsuperscript{214} DUBERSTEIN, M.J, pp.122-123, see note 211.
\textsuperscript{215} Article XX NFL CBA.
\textsuperscript{216} The definition of what constitutes salary is somehow complicated (see paragraph 2(c) of Article XIV of the NFL CBA for the definition), but it is not necessary to analyze it in order to make my argument.
resulting from the use of this restriction would be equal to the one that the player would be able to get on the free market.

For teams, the benefits of using these designations are limited since they are only are able to use each one on one player and for only one year. Furthermore, multiple uses across the league will have an inflationary effect increasing the average salary of the players at that position so the high salaries resulting from the use of that instrument induce the teams to use it with caution.

The influence of the designations on the value of players is the same as the one described previously in the restricted free agent section\textsuperscript{217}. But in this case, the reduction of salary may be less than for a restricted free agent since it makes the player one of the best paid players in the league. Again we are faced with an instrument designed to artificially lower the salaries of a player in order to allow the teams to capture part of the value of the player for themselves.

**5.4. Entering the players pool, the college draft**

In this section, I will examine one of the most unique features of North American sports, the draft. This way of distributing young players remain the most important restriction on the free movement of players in these leagues. As a representative example, I will analyze the NFL draft in details. This will allow drawing conclusions that are valid for all other North American sports since they use similar drafts to allocate young players between the different teams. Analyzing the NFL draft in principle will also allow us to understand the general idea behind that mechanism.

**5.4.1. Training and development of young players in football**

In football, the development of young players is independent from the professional teams of the NFL\textsuperscript{218}. The structure of amateur football makes the academic institutions, colleges and high schools, responsible for the development of players. These amateur teams are independent from the NFL and compete in their own closed leagues with different eligibility

\textsuperscript{217} Section 5.3.3.2.

\textsuperscript{218} In 1991 the NFL started operating the NFL Europe with the double goal of creating a European market for football and the development of young players. The teams of that league were composed of players of the NFL teams assigned to the European teams to further up their formation. However, at the end of the 2007 the NFL teams decided to end the operation of that league.
rules for the players. The level that interests us for this section is the college level in which players must be at least 18 years of age and are allowed to play up to six years for these teams. When a player leaves for the NFL, the college team receives no compensation. All the investments made in the training and developments of the players are lost. College teams knowingly get “free ridden” on by the NFL, developing the players and letting them go when they reach their prime. These teams are willing participants in the system and do not ask for any form of compensation. They also have incentives to develop players in order to have success in their own championships. Furthermore, players at that level are not remunerated, remuneration being forbidden and that rule is being strictly enforced with stiff penalty to colleges that do not respect it. This allows the teams to capture all the revenues of the sport. College football being a 2 milliard $ (US) industry that benefits from a 25% growth rate\textsuperscript{219}, the loss of the best players is compensated by the fact that they are forced to play for free for for-profit organizations for a minimum of three years\textsuperscript{220}. Since the object of this work is not to analyze College football, what is important to remember for my purpose is that future NFL players get free ridden on by the College teams that employ them for free and capture all their value. Therefore, it is only just that they are allowed to leave the system for free. We have now seen the reasons why the lower echelons of football allow their players to enter the draft. In the following sections, I will look at the draft itself, its structure and its legal and sporting consequences.

5.4.2. The NFL college draft

The draft consists of seven rounds in which each team is allowed to select one player. When drafting a player, the team gets the exclusive rights to negotiate a contract with this player since it makes him a rookie non-free agent. The selection is usually done in the reverse order of the results of the last year championship with the worst team having the first choice. The logic behind that method is to allow the worst teams to improve through the addition of the best young players. This should in theory create a more balanced league since the arrival of new talent will be concentrated in the teams that had the poorest results. But from the point of view of the players, the draft mainly looks like a system to restrain their ability to offer their services to the team that would value them the most. In consequence, they will receive lower salaries when they enter the league. In the following sections, I will first address the latest

\textsuperscript{219} \textit{FORBES}, 2008. The Business of College Football, 13\textsuperscript{th} August.

\textsuperscript{220} More on that in section 5.4.2.
attempt made by a player to contest the NFL draft. Secondly, I will examine if the draft helps the NFL maintain a better competitive balance between its teams and finally, I will discuss the fact that the draft may not only be to the benefit of the NFL teams and that some players may actually get higher salaries because of the particular negotiation situation that the draft imposes on the implicated parties.

5.4.3. Contesting the draft

In 2004 the NFL draft was contested by a future player, Maurice Clarett\textsuperscript{221}. The NFL rules require that a player, who wants to enter the college draft, should have graduated from high school, the academic institution before college, for at least three football seasons\textsuperscript{222}. After a very successful first season Mr. Clarett was suspended from playing at the college level for various academic infractions. This suspension made him unavailable to play for the whole college football season. After not playing for one complete season, Mr. Clarett decided to enter the 2004 NFL college draft. He was refused the right to be a candidate by the NFL since he did not meet the criteria to enter the selection, having graduated from high school only two football seasons earlier.

Mr. Clarett then sued the NFL arguing that the criteria required to enter the draft were an unreasonable restriction to the entry to the market for professional football players and that it constituted a violation of the U.S. antitrust laws. He asked for an injunction allowing him to enter the draft. He was successful at the first instance, the judge issuing an injunction allowing Mr. Clarett to enter the draft, but the NFL appealed. The Appeal Court reversed the initial judgment and rejected Clarett’s claim. The Court said that the restrictions to the entry of the draft were part of a collective bargaining agreement and that this was protected from the antitrust rules by the non-statutory labour exemption\textsuperscript{223}. The Court also said that even if the draft eligibility rules were not explicitly written in the NFL CBA, they were known to the players’ union which had the opportunity to negotiate if it wanted to change them. The US Supreme Court refused to hear the appeal of Mr. Clarett on the injunction, so the rules of the NFL College draft survived.

\textsuperscript{221} Clarett \textit{v.} NFL, United States Court of appeals for the 2\textsuperscript{nd} Circuit, docket 04-0943, May 24 2004.

\textsuperscript{222} Article XII of the NFL Constitution and Bylaws.

\textsuperscript{223} McCoy, see note 38 and Radovich see note 9.
However, the NFL did not seem to be optimistic about its chances of winning if the merits of the case had been heard by the Supreme Court since it negotiated the addition of the following section in the July 3 2006 extension of the CBA:

“(b) No player shall be permitted to apply for special eligibility for selection in the Draft, or otherwise be eligible for the Draft, until three NFL regular seasons have begun and ended following either his graduation from high school or graduation of the class with which he entered high school, whichever is earlier. For example, if a player graduated from high school in December 2006, he would not be permitted to apply for special eligibility, and would not otherwise be eligible for selection, until the 2010 Draft.

(…)

(d) No player shall be eligible to be employed by an NFL Team until he has been eligible for selection in an NFL Draft.”

By doing so, the NFL strengthened the argument that the draft is covered by the labor exemption but this has not been tested yet.

5.4.4. Empirical evidence regarding the influence of the draft on competitive balance in the NFL

The draft may now be safe from an anti-trust contestation but does that make it achieve its goal? In the Clarett decision, there was no discussion about the goal of the draft, the only question that interested the court was to know if antitrust law applied to it or not. So we have to go in an empirical analysis of the draft to find out for ourselves if the draft attains the goal of helping the bad teams to get better.

An interesting study of the use of draft picks by the teams of the NFL was done by sports writer Paul Zimmerman. Although there are some imperfections such as the fact that it covered only a ten years period (1994 to 2003), that only the first three rounds of the draft (out of seven) were analyzed and that some of the data used is subjective (like the definition of the term regular and flunk), it is still useful in order to understand the impact of the draft on the individual NFL teams.

224 Section XVI (2) of the NFL CBA 2006 Extension
The idea behind Mr. Zimmerman’s exercise was to try to find the retention rate of the NFL teams of the high draft pick players. Mr. Zimmerman decided to take only the first three rounds in order to separate the impact players from the average ones since a player drafted in the first three rounds should be able to have enough talent to have a decent career in the NFL.

I used the original data that can be found in Mr. Zimmerman’s article in order to create the following table:

<table>
<thead>
<tr>
<th>TEAM</th>
<th>REC</th>
<th>REM/TOT</th>
<th>REG</th>
<th>REG %</th>
<th>F</th>
<th>W/L</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arizona</td>
<td>(57-103)</td>
<td>15 of 34</td>
<td>17</td>
<td>50</td>
<td>2</td>
<td>L</td>
</tr>
<tr>
<td>Atlanta</td>
<td>(70-89-1)</td>
<td>7 of 24</td>
<td>11</td>
<td>46</td>
<td>5</td>
<td>L</td>
</tr>
<tr>
<td>Baltimore</td>
<td>(79-80-1)</td>
<td>14 of 26</td>
<td>17</td>
<td>65</td>
<td>2</td>
<td>L</td>
</tr>
<tr>
<td>Buffalo</td>
<td>(79-81)</td>
<td>14 of 35</td>
<td>18</td>
<td>51</td>
<td>3</td>
<td>L</td>
</tr>
<tr>
<td>Carolina</td>
<td>(64-80)</td>
<td>14 of 28</td>
<td>16</td>
<td>57</td>
<td>2</td>
<td>L</td>
</tr>
<tr>
<td>Chicago</td>
<td>(68-92)</td>
<td>15 of 32</td>
<td>17</td>
<td>53</td>
<td>3</td>
<td>L</td>
</tr>
<tr>
<td>Cincinnati</td>
<td>(52-108)</td>
<td>10 of 32</td>
<td>18</td>
<td>56</td>
<td>5</td>
<td>L</td>
</tr>
<tr>
<td>Cleveland</td>
<td>(26-54)</td>
<td>12 of 18</td>
<td>9</td>
<td>50</td>
<td>1</td>
<td>L</td>
</tr>
<tr>
<td>Dallas</td>
<td>(83-77)</td>
<td>14 of 31</td>
<td>17</td>
<td>54</td>
<td>3</td>
<td>W</td>
</tr>
<tr>
<td>Denver</td>
<td>(98-62)</td>
<td>11 of 27</td>
<td>12</td>
<td>44</td>
<td>3</td>
<td>W</td>
</tr>
<tr>
<td>Detroit</td>
<td>(65-95)</td>
<td>11 of 29</td>
<td>19</td>
<td>66</td>
<td>1</td>
<td>L</td>
</tr>
<tr>
<td>Green Bay</td>
<td>(108-52)</td>
<td>14 of 31</td>
<td>16</td>
<td>52</td>
<td>2</td>
<td>L</td>
</tr>
<tr>
<td>Indianapolis</td>
<td>(83-77)</td>
<td>14 of 31</td>
<td>20</td>
<td>65</td>
<td>1</td>
<td>W</td>
</tr>
<tr>
<td>Jacksonville</td>
<td>(73-71)</td>
<td>12 of 32</td>
<td>22</td>
<td>69</td>
<td>2</td>
<td>W</td>
</tr>
<tr>
<td>Kansas City</td>
<td>(94-66)</td>
<td>11 of 28</td>
<td>14</td>
<td>50</td>
<td>0</td>
<td>W</td>
</tr>
<tr>
<td>Miami</td>
<td>(96-64)</td>
<td>12 of 32</td>
<td>16</td>
<td>50</td>
<td>7</td>
<td>W</td>
</tr>
<tr>
<td>Minnesota</td>
<td>(92-68)</td>
<td>13 of 32</td>
<td>19</td>
<td>59</td>
<td>3</td>
<td>W</td>
</tr>
<tr>
<td>New England</td>
<td>(93-67)</td>
<td>13 of 35</td>
<td>17</td>
<td>49</td>
<td>3</td>
<td>W</td>
</tr>
<tr>
<td>New Orleans</td>
<td>(66-94)</td>
<td>11 of 27</td>
<td>18</td>
<td>67</td>
<td>0</td>
<td>L</td>
</tr>
<tr>
<td>NY Giants</td>
<td>(78-81-1)</td>
<td>13 of 31</td>
<td>17</td>
<td>55</td>
<td>0</td>
<td>L</td>
</tr>
<tr>
<td>NY Jets</td>
<td>(73-87)</td>
<td>14 of 31</td>
<td>17</td>
<td>55</td>
<td>1</td>
<td>L</td>
</tr>
<tr>
<td>Oakland</td>
<td>(81-79)</td>
<td>17 of 31</td>
<td>17</td>
<td>55</td>
<td>2</td>
<td>W</td>
</tr>
<tr>
<td>Philadelphia</td>
<td>(87-72-1)</td>
<td>15 of 36</td>
<td>18</td>
<td>50</td>
<td>2</td>
<td>W</td>
</tr>
<tr>
<td>Pittsburgh</td>
<td>(95-64-1)</td>
<td>14 of 34</td>
<td>17</td>
<td>50</td>
<td>2</td>
<td>W</td>
</tr>
<tr>
<td>St. Louis</td>
<td>(82-78)</td>
<td>15 of 39</td>
<td>20</td>
<td>51</td>
<td>6</td>
<td>W</td>
</tr>
<tr>
<td>San Diego</td>
<td>(63-97)</td>
<td>11 of 33</td>
<td>14</td>
<td>42</td>
<td>3</td>
<td>L</td>
</tr>
<tr>
<td>San Francisco</td>
<td>(100-60)</td>
<td>14 of 31</td>
<td>15</td>
<td>48</td>
<td>3</td>
<td>W</td>
</tr>
<tr>
<td>Seattle</td>
<td>(78-82)</td>
<td>16 of 32</td>
<td>16</td>
<td>50</td>
<td>1</td>
<td>L</td>
</tr>
<tr>
<td>Tampa Bay</td>
<td>(86-74)</td>
<td>12 of 28</td>
<td>21</td>
<td>75</td>
<td>2</td>
<td>W</td>
</tr>
<tr>
<td>Tennessee</td>
<td>(89-71)</td>
<td>14 of 33</td>
<td>18</td>
<td>55</td>
<td>2</td>
<td>W</td>
</tr>
<tr>
<td>Washington</td>
<td>(70-89-1)</td>
<td>11 of 26</td>
<td>15</td>
<td>58</td>
<td>2</td>
<td>L</td>
</tr>
</tbody>
</table>

REC: Playing record of the team over the ten years (wins, losses, ties)
REM/TOT: Players still with the team, from the total number drafted in the first three rounds of 1994 through 2003
REG: Regulars who developed from that group, a regular being defined as a player that played at least 50% of the games in the studied period, for the team that drafted him or another NFL team
REG %: Percentage of the players drafted by the team that developed into regulars, with the team that drafted him or with another team
F: Flunks; players who lasted one year or less in the league before being fired
W/L: Winning or losing record of the team in the years studied

The five teams with the best percentage of Regulars are highlighted in yellow and the worst five put in boxes.

As one can see from the data, the losing teams do not seem to have a better retention rate of players than the winning ones. Although this study does not control for a variety of factors (like scouting budget and injuries to players), I believe that it offers empirical support to the theory that the draft does not significantly affect the competitive balance of the NFL. The team with the best pick/regular ratio in the studied period is Tampa Bay with 75% and a winning record. This means that the team would not have had the opportunity to draft the most talented players since the winning record means that it would pick late in the draft. The second best team (Jacksonville) is also a winning team and we have to go down to the 3rd and 4th teams (Detroit and New Orleans) to find losing records. At the other end of the spectrum, the teams with the worst percentage of regulars are two losing teams (San Diego and Atlanta) and one winning team (Denver).

This data shows that the draft does not help the bad teams to change their records. Seeing the four top teams with the best percentage of regulars, only one has a record that could suggest improvements (Jacksonville with a 73-71 record). The other three have pretty solid good or bad records, so the drafts most likely did not influence their performances during the ten years studied.

Additionally, the definition of regular does not only look at players who remained with the team that drafted them, therefore it is not directly correlate to the actual improvement of the team but to the potential improvement by showing the talent available to the team in the draft. It controls for the error in judgment and the weakness of the development system of a team by considering players that reach their full potential with other teams. If the player flourishes
with another team that still means that the team that drafted him had the potential to improve itself.

So the rank in the draft is not the only way to predict the performance of a new player in the league and other factors like scouting, development budget and the skills of the coaches may have a greater influence on predicting the performances of a new player in the league

In fact, over the years the structure of the draft has evolved in a way that also strengthens the negotiating power of the players. In some cases a win draft position may even be a winner’s curse.

5.4.5. The draft, a winner’s curse?

There may be another reason why the draft fails to have an effect on the competitive balance of the league. An analysis of the various players selected in the draft, their costs and their impact on the on-field performance of teams seems to confirm that selecting a player early in the draft is not a guarantee of improving the team. It was found in the same study that the draft method of the NFL is inefficient, at least in the first round, which is where the best players are selected. According to this study, the players drafted early are able to exert a monopsonic rent from the teams that select them because of their high draft pick status. This situation has the potential of turning into a winner’s curse for the team since the on-field production of the player will not be justified by his high salary. Even if players can only negotiate with one team, the team on its behalf can only negotiate with the players that it has drafted. This monopsony-monopoly relationship results in teams overpaying their first round draft picks, either because of their dire need for talent or because of the pressure not to lose face by not being able to sign the player that they have selected. This study also showed that it is around the 43rd draft pick that a team can find the best value for its money. Players selected at this moment provide an on-the-field production that exceeds their salary, allowing the team to “make a profit” in the sense that was explained in the equation in chapter 3. After that, the chances of drafting a productive player reduces dramatically, even if the players are signed at a low cost since they have a smaller chance of contributing to the team’s success.

---

228 Section 3.5.3
The 43rd draft pick corresponds to the middle of the 2nd round and is usually attributed to teams that had an average performance in the last season. Picking an effective player at this point is more related to the scouting expertise of the team and the competence of its managers than its position in the championship. From that perspective the draft seems designed to restrict the player’s movement in the first years of his contract and not to make sure that bad teams get better with time. If all rookie players were allowed to negotiate with all teams, they would be able to have a salary that corresponds to their real market value. But for some players the draft is a benefit since it allows them to get a “draft order bonus” that enables them to get a salary that is higher than what the market would be willing to pay for their services. As for the teams, having the possibility to sign any available player would give them the incentive to make a better cost benefit analysis and to go after the players that they actually need to improve themselves and offer the players a salary that is related to their talent.

If we look at the draft from a competitive balance point of view, it is not necessary in modern professional sports since it does not seem to have any impact on them. The only purpose that it appears to serve today is to distort the market for the service of professional sport players. It could be argued that the draft is now unnecessary and that a professional league with a salary cap could allow the new players to negotiate their entry in the league with the highest bidder. This would allow more competitive balance since the worst teams could pick up any amount of promising new players as long as they respect the salary cap. They would therefore be able to improve faster than under the current regime, in which they only get to pick one or two potential good players, allowing the league to be more competitive and improving its value by creating more demands for games. This would be a win-win situation for the players and teams since salary caps are linked to the league’s total revenues. This may not make the team automatically smarter and take better scouting decisions, but it would give them the incentives to do so.

230 MASSEY, C. AND THALER, R., see note 227.
5.4.6. What is the draft really about?

Based on the data presented in the past sections, the fact that a team has an opportunity to select players early in the draft does not seem to help it improve the team’s performance on the field.

It seems that the draft serves two purposes. First, as Mr. Clarett argued in his case\(^{232}\), it is a way of restricting the freedom of movement of the players and in consequence a lowering of their salaries by limiting their negotiation strength. Second, drafts have traditionally been explained by North American sports leagues by the fact that they are needed to help the teams finance the development system in which young players are prepared for the big leagues\(^ {233}\). By paying the players a salary below their market value, the teams are able to invest the savings in the training and development of these players. So in theory the teams do not appropriate any of the value of the players. This reasoning is not applicable to the NFL since it receives the players at an older age and they have already been trained and developed by their College teams. But the situation of the NFL is unique; all other North American sports have to invest in an extensive development system for their young players\(^ {234}\). Therefore this could be a valid argument in favor of the teams and show that the draft is more than just an instrument used by the teams to capture part of the value of the players for themselves. But this remains hard to prove.

5.5. The salary cap

As mentioned earlier, the teams in the NFL operate under the constraint of a salary cap. This means that all teams cannot spend more than a certain amount of money in a given year on the salary of their players. This arrangement creates in fact a partnership between both teams and players since they agree in advance on the distribution of revenues of the business between themselves. Additionally, the cap is linked to a floor that forces teams to spend a minimum amount of money on players. So it would be more appropriate to talk about a salary spread instead of cap. It could be seen as a simple way of reducing salaries of players, but it actually

\(^{232}\) Clarett v. NFL, see note 221.


\(^{234}\) Interview with J. Brisebois, see note 114 and + KRAUTMANN., A. AND OPPENHEIMER, M., see note 233.
makes the players the stakeholders of a certain percentage of the business of the NFL and guarantees them a certain level of revenues.

5.6. Under the shadow of the law

We can see in the provisions governing the movements of players in the NFL an example of negotiation “under the shadow of the law”. The McNeil decision was indeed a great victory for the players, but with the court’s recognition of the right of the league to implement measures to restrain the movement of players, the players knew that they would never be able to get total freedom of movement. Following the decision, both parties had incentives to negotiate a new set of rules.

Looking at the current system, it is hard to see how it could be attacked by a future antitrust action since it would probably be protected under the non-statutory labour exemption. The only danger would be if the CBA came to an end and the restriction system would be unilaterally imposed by the teams. However, the parties put incentives in the CBA to make sure that it would always be renegotiated a long time in advance and in order to prevent a conflict. The CBA makes sure that when it runs into its last year, both parties will end up losing; the teams by losing the protection of the salary cap and the players by having to wait one more year in order to become unrestricted free agents\textsuperscript{235}. Therefore, rational parties would never let the CBA reach its final year.

It is very clear what the teams and the union attempted by explicitly including the draft in the 2006 CBA extension. They wanted to make sure that the draft was from then on officially covered by the labour anti-trust exemption so that there are no chances of facing a second challenge like the one made by Mr. Clarett. This exemption is the only protection for the draft mechanism since the “scientific” argument put forward by the NFL for that restriction saying that the players are not physically and mentally ready to enter the draft before around 21 years of age\textsuperscript{236} does not have a lot of credibility. In 2007, a 19 year old player was exceptionally allowed to enter the draft since he had graduated from high school at the age of


\textsuperscript{236} \textit{Clarett v. NFL}, see note 221.
16. Even if he should theoretically be considered not ready to play in the NFL, he was drafted and performed well from the start.  

5.7. Effect of these measures on the value of players

The negotiated solution that is contained in the NFL CBA does not only affect the freedom of movement of the players, but it also casts a long shadow on the whole determination of value of the players. In the following sections, I will examine the general impact of the CBA on the value of players, the impact of free agency and limitation on players’ remuneration and finally the impact of salary caps on the value of players.

5.7.1. The impact of a CBA on the value of a player

It is one of the great ironies of the sport world that the USA, one of the countries that are the least friendly to labour unions, is leading the way in unionisation of professional sports players. The presence of a CBA allows players to have a certain amount of contractual conditions negotiated by the union and it allows them to concentrate their negotiation on their individual working conditions, knowing that some aspects of their contract has already been negotiated by the NFLPA. Professional sport CBAs regulate some important aspects of the relationship between the players and the teams - from the number of players allowed per team to the disciplinary sanctions that can be imposed on a player. This resulted in the league and the union agreeing on a standardized playing contract that states the main working conditions of the player without him having to individually negotiate them. The only thing that is not covered is the players’ individual remuneration. There is a minimum salary for players stated in the CBA, but the individual amount of compensation is left to the player and his team to negotiate. Essentially, both parties negotiate on the numbers to put down on the line of a standardized form contract in which 90% of the conditions have been pre-negotiated by the union and the teams. The players benefit from the best of both worlds: on the one side the strength of the collective bargaining process for their general working conditions and on the other their individual achievements for their salary without the interference of a standard pay scale.

---

So concerning the players, the CBA allows them to concentrate on trying to get their maximal individual value in the form of remuneration without having to worry about the non-monetary issues since these have already been negotiated by the union. So the negotiations crystallise the value of the player in a monetary form as non-monetary compensation is uniform for all players and not the object of negotiation. These standard conditions are known by both parties at the start of their negotiations and allow them to attribute value to them from the start. They can afterwards use this value to determine the salary that is due to the player in accordance to the formula that was determined in chapter 3. With this hybrid arrangement the NFL players try to capture the greatest value for their skills in the form of a salary and general non-monetary advantages and the object of the negotiation is to determine the distribution of the value between the player and the team with part of it already agreed in the form of non-monetary compensation.

5.7.2. The impact of free agency on the value of players

The NFL judicial saga was at the forefront of the battle of professional players to be able to benefit from a competitive market for their services. But what was the impact of free-agency on the value of players? Did they suddenly become more valuable because other teams could bid for his service?

It is highly unlikely that more freedom for a small percentage of players created new value. Assuming that no restrictions on the movement of players exist, only the players at the end of their contract would benefit, in a given year, from the free market for their services. The effect of a competitive market for the players at the end of their contracts does not create new value; it is more likely to result in a redistribution of the monopolistic rent of the business of sport in favour of these players. The monopolistic rent in sport has traditionally been distributed between the teams, via high profits, and the fans, via low ticket and merchandising prices. This was done at the expense of the players that had to suffer restrictions on their movement resulting in lower salaries. By removing these barriers, it shifted the distribution away from the fans and the team in favour of the players. The fans have to pay more for tickets and merchandising and the teams have to satisfy themselves with lower profits. By being free to offer their services to the highest bidder, the players are able to get a salary that is close to

---

238 DOBSON, S. AND GODDARD, J., see note 47.
their economic value. To refer again to the formula in chapter 3, $S$ grows and forces $ES$ close or equal to 0.

However, respecting the sportsman owner effect\textsuperscript{239}, there will be a tendency for teams to overpay for the services of players. As stated earlier, this will lead to an inflation of the salaries and a reduction of the value of the team since it can drive $ES$ to negative values. This has the potential to have negative impact on the competitive balance of the league and in consequence on its overall revenues. Therefore, a rational league and player union will try to cooperate and create a partnership to prevent that and use mechanisms like the salary cap to rein in Sportsmen owners and insure the prosperity of the league.

But the league and the union do not seem to see the restrictions to the movement of players in that perspective. It is generally argued that these restrictions are in place to help the competitive balance of the sport. Although this may be true indirectly, by allowing all teams to survive and limiting the impact of the sportsman owner, from the value oriented point of view of this thesis, the result is clear: the free movement of players allows them to appropriate more and more of the monopolistic rent of sport for themselves. The restrictions of the players’ freedom of movement result in a greater part of the rent remaining with the teams. It has little to do with competitive balance and more to do with revenue and monopolistic rent distribution. The introduction of restrictions on the movement of players influences their value - not the individual value of each player but their aggregate value - by restricting the amount of the monopolistic rent of sport that is allocated to them via their salaries.

5.7.3. The impact of salary caps on the value of players

In the case of a salary cap, the parties have agreed on the amount of the monopolistic rent that is going to be allocated to the players in the form of a percentage of the league’s total revenues. A salary cap is, in that perspective, a contract between the teams and the players on how to split the rent of the sport. They become business partners and determine together the aggregated value of the players. Salary caps also have a second effect, the reduction of the players’ salaries\textsuperscript{240}. Salary caps keep the sportsmen owners in check by putting a limit to the amount of money that they can give to their players, a limit that does not exist in an uncapped

\textsuperscript{239} Section 3.3.
\textsuperscript{240} KÈSENNE, S., see note 20, AHLBURG, D. AND DWORKIN, J., see note 182 and GRAVES, J., see note 231.
league. So if the players get lower salaries by agreeing to a salary cap, why do they agree to it after all?

First, they hope that by having a league that is financially healthy, the revenues will grow and all players will benefit from it in the end. In short, they bet that the “pie” will get bigger instead of fighting for the bigger slice of the pie.

Second, even if the working life of a professional athlete is relatively short\(^2\), the unions that negotiate these agreements have been in place since a long time and will still be in place decades in the future. They are, when agreeing to caps, taking the long term view that a healthy sport will allow the revenues to grow and the players will benefit from it in the end. By limiting the expenses all teams will hopefully be able to be competitive in the long term and create more uncertainty of results in the league. This should also result in more revenues and greater financial stability for the league and the individual teams through a higher demand for the sport.

What the salary cap represents is an agreed limit of the aggregated value of the players. The hope of the parties is that it will make the whole business more valuable and help generate extra growth that will be distributed between the teams and the players\(^2\). Therefore, the introduction of a salary cap is an agreement between the teams and the players on the aggregate value of the players instead of letting the free market determine it.

5.8. The impact of a change in the commercial rules of the game in football on the value of players

The football rules of the commercial game have been extensively negotiated by the players and the teams. It seems, after years of vigorously contested legal actions, that the parties realised that they were better off negotiating the commercial rules of the game than letting the


courts decide. They understood that they were partners in the business of football and that they had to collaborate in order to make sure that the business runs smoothly. With restrictions on the movement of players and a salary cap included in a CBA, the players became the business partners of the teams and had incentives to make sure that the sport be financially successful. Both sides would from then on benefit if the sport grew. But the influence of a CBA on the growth of a sport is limited. Growth in sport is most likely to come from other factors like a growing gross national product, the enrichment of the middle class, the influence of radio and television that create new sources of revenues and the expansion of the league in new geographic markets. With the CBA, the only influence possible is exerted by helping the competitive balance of the sport and in consequence helping the demand for football grow. The actual trend in North American sports makes the CBA an important agent of redistribution of the monopolistic rent of sport between the parties. The players and the owners sit together to determine the collective value of the players in advance. By way of negotiation, they came up with a set of rules that directly or indirectly crystallised this value. Teams and players hope to build a business model that will stimulate the growth of the sport by negotiating the distribution of the rent. With growth should come higher revenues and result in higher value for both parties. A CBA helping the growth of the business of sport, or at least the growth of monopolistic rent of sport and including instruments to limit the market for the services of players could therefore be seen as a mean to influence the value of players. Here we have an example where the commercial rules of the game reflect the agreed upon value of the players, in the hope that it will result in more growth to the benefit of both parties. This indirectly affects individual players, giving them a bigger pie to claim money from, but not the size of their individual servings. The commercial rules of the game of football do not affect directly the individual players; they have, as we have seen in that section, an important effect on it.

5.9. Soccer

In the case of soccer, the sport was allowed to operate almost as it pleased until 1995. The Fédération International de Football Association (FIFA), the international organisation regulating soccer around the world, escaped specific countries’ legislation because of the fragmented nature of soccer and its international appeal. Until 1995, freedom of movement of soccer players was restricted on two levels. First, there was a reserve clause that made a player the property of the team that signed him, even if the employment contract had expired.
Second, the various organisations that ruled soccer in each country, with the blessing of FIFA, limited the number of foreign players who were allowed to play in their national championship. These rules were successfully challenged in the now legendary *Bosman* case that opened up the market for the services of soccer players. The first part of this section will look at the case, its consequences and the new transfer rules that were adopted following it. The second part will examine how these new rules impacted the value of soccer players.

### 5.9.1. The *Bosman* case

A lot has been written on the *Bosman* decision. For the purpose of this work, I will limit my analysis to its impact on the soccer players by examining the changes that it brought to the market for their services. To make an effective analysis, we need to understand the facts that led to the decision.

#### 5.9.1.1. History of the case

Jean-Marc Bosman was a promising young Belgian soccer player. At the expiration of his contract he ended up in a dispute with his team. Believing that a negotiated solution was impossible, he asked to be transferred to another team. A deal was made to transfer Mr. Bosman to a 2\textsuperscript{nd} division French team. However, the old team had doubts about the financial stability of the French team and backed off from the transfer at the last minute. Unsatisfied with the offers made by his current team and the missed transfer opportunity, Mr. Bosman refused to report to the team. This resulted in a situation in which it was impossible for Mr. Bosman to earn a living since on the one hand he was not able to agree on the terms of a new contract with his current team that had the exclusive rights to his services and on the other he could not play in Belgium or anywhere else since he refused to report to his teams and his employer refused to transfer him. Mr. Bosman therefore sued, asking for a permanent injunction allowing him to continue his career and compensation for his loss of revenues. His main argument was that the transfer system was a violation of the right of free movement of persons within the European Union (EU) and a violation of the EU competition laws. The matter was referred to the European Court of Justice (ECJ) by the Belgian court. The ECJ

---

\(^{243}\) *Bosman*, see note 9.

\(^{244}\) The judicial war between Bosman, his team and the Belgian soccer union gave rise to multiple actions, motions and appeals that are not all mentioned here. The goal of the following summary is to give the reader a brief outlook of this saga and in doing that it is not appropriate to examine every battle that was fought.
had to determine the legality of the soccer transfer system. Even if the transfer system that was the object of the lawsuit was only the Belgian one, there was an implication for the international system. Indeed, the Belgian transfer system had been approved by FIFA and was similar to the transfer mechanisms in most of the other major soccer countries. The ECJ decided that the transfer fees charged for a player who had ended his contractual relationship with the team were an illegal restriction of the free movement of persons within the EU and in consequence illegal and unenforceable in the EU. Regarding the competition law argument and the possible restriction of the market for players, the ECJ refused to consider it since it was unnecessary for its ruling. Using freedom of movement was the simplest solution and the ECJ probably had the impression that its decision would give the best results without having to address complex competition law questions. So with that ruling, the ECJ judged that the transfer system that was used in soccer since the end of the XIX\textsuperscript{th} century was illegal and that the way the industry operated its business needed to be changed.

5.9.1.2. What are transfer fees?

This question may seem superfluous for the soccer fan, but for the reader who does not necessarily know the ins and outs of this sport, a short explanation is needed. At the time of the \textit{Bosman} decision there were two ways for a soccer team to get the rights on a player. First, it could train him from the beginning via academies or related amateur soccer teams affiliated with the professional team. The second way was to “buy” the rights to field a player from another team which is called a transfer. The money paid from the buying team to the selling team is called a transfer fee. The North American sports fans are familiar with the concept of trades where teams trade players for other players. However, the tradition in soccer is to transfer players for money.

5.9.1.3. The pre-\textit{Bosman} transfer system

Under the pre-\textit{Bosman} system, teams would get a fee every time they transferred a player to another team irrespective of his contractual situation. There was always a certain value

---

\textsuperscript{245} \textit{Bosman}, paragraphs 138, see note 9.

\textsuperscript{246} Even in North America the transfer of players for money has been allowed and used in the past. The most infamous example being the sale of legendary player George Herman “Babe” Ruth to the New-York Yankees by the Boston Red Sox in 1920. After the sale, Ruth became one of the best baseball offensive players in the history of the game and the Red Sox went 80 years without winning a championship.
attached to the players so teams could not lose them for nothing when the player’s contract had expired. However, this could lead to situations in which the price asked for a player would be higher than what the other teams were willing to pay for him. The player would then suffer having to accept a new unilateral contractual offer of his team that usually included a sharp decrease in revenues or be willing to go on strike until he was able to get transferred. This is exactly what happened to Mr. Bosman. To use the formula from chapter 3247, the situation allowed the team to raise AC to levels that would make the player unattractive to any other team since this high AC would result in a negative ES when combined with the player’s salary (S).

This was not a new problem in 1995 since there are previous English cases248 in which players had sued their teams to get transferred. The first one dates back to 1912249. *Eastham vs. Newcastle United* was the first successful challenge of the English transfer rules that were at the time more restrictive than the rules invalided in the *Bosman case*. In addition to the restriction on the free movement of players, teams had introduced a maximum salary that capped the amount a player could receive. In Mr. Eastham’s case, the court ruled that the transfer fees affecting him were

“(N)ot binding on the plaintiff and are unreasonable restraints of trade”250.

So thirty years before the *Bosman* case, freedom had already been granted to the English soccer players. This did not have more of an impact on the soccer world for the following reasons: first, there was at the time no strong players’ association. This prevented the players from using the decision as leverage during the negotiation of new rules; new rules were drafted but they only gave extra advantages to the good players and almost none to the marginal players.251 Teams used the lack of solidarity between the players to make sure to maintain control on the movement of players. In essence, the post-*Eastham* system still prevented the free movement of players without a valid employment contract. The second reason *Eastham* had little impact is that it was hard to reproduce outside of England and

---

247 Section 3.5.3.
249 *Kingaby v. Aston Villa* (unreported)
250 *Eastham v. Newcastle United*, paragraph 160, see note 65.
251 GREENFIELD, S., see note 65.
Wales. Unlike *Bosman*, the decision was not based on EU Community Law but on the Common Law contract doctrine of “restraints of trade”. This judicial concept was hard to transfer to the continental Europe civil law jurisdictions that lacked a comparable legal instrument. So even if players did actually have the opportunity of attacking the transfer system, they failed to capitalize on it.

5.9.1.4. After *Bosman*

The *Bosman* case was strongly denounced by the soccer authorities and the end of soccer as we know it was forecasted. Such apocalyptic declarations seemed more motivated by a political aim than by reality. One must remember that the ruling of the ECJ concerned at the time only about ten percent of the active soccer players in Europe, the ones at the end of their contracts. They could then benefit from a strengthened negotiation position and “test the market” without having to worry about the transfer fee that would be asked for their services by their old team. The rest of the players were still affected by the transfer rules and did not gain more freedom since they were still bound to their team by their playing contract. It is interesting to note that despite the bleak predictions, business returned pretty much to normal for soccer after *Bosman* and there was no destruction of the industry. Players and teams adjusted their behaviours in order to make the best out of the situation. Teams that wanted monetary compensation if a player left transferred him before his contract ended and received a transfer fee instead of receiving on-the-field performance and the associated revenues for the remainder of his contract. Additionally, the teams signed their valuable players to long term contracts. This behaviour ended up being exactly as investment theory and the Coase theorem would predict with each team trying to maximize their return on investment and each player ending up at the service of the team that valued him the most. Each party therefore adjusted its behaviour in order to maximize revenues and performances in the new system. So the forecasted apocalypse did not arrive and soccer survived, even if some still think that it is only a question of time before the whole system will collapse.

---

252 SZYMANSKI, S. AND ZIMBALIST, A., see note 57.
254 Idem.
255 Idem.
5.9.2. The birth of the new system

Following the *Bosman* decision, FIFA started to design new rules to regulate the transfer of players. These rules were the object of negotiations with the European Commission in order to insure their compliance with EU law. The Competition Commissioner at the time, Mario Monti, was reported to have been personally implicated in the drafting of the rules, so most observers call them the Monti rules.

Although it invalidated the old transfer system, the ECJ, in the *Bosman* decision, left the door open for FIFA to draft regulations to restrict the free movement of players, but under certain conditions. The Court wrote, in paragraph 106

“In view of the considerable social importance of sporting activities and in particular (soccer) in the Community, the aims of maintaining a balance between the teams by preserving a certain degree of equality and uncertainty as to results and of encouraging the recruitment and training of young players must be accepted as legitimate.”

(My underlining)

In that statement, the ECJ indirectly says that it would be possible for the teams, under certain conditions, to behave contrary to the European Laws. It is therefore crucial, before analyzing the transfer rules, to look at that jurisprudence test.

5.9.3. The *Bosman* test

Of all the exceptions the ECJ could have used, it chose the traditional economic justifications for the transfer fees.\(^{257}\) Traditional economic theory of the transfer fees in soccer states that with this mechanism, the richer teams are prevented from buying all the good players and all teams are given incentives to recruit and train young players. By having to pay a transfer fee, the cost of acquiring the player is higher and it should prevent the richer teams from “buying” a championship team every year. Also, when such a sale happens, the money received allows the selling team to pay for new players or to give better salaries to the remaining ones. This should, in theory, result in a more equal repartition of soccer talent between the teams.

\(^{257}\) DOBSON, S. AND GODDARD, J., see note 47.
The presence of a transfer fee system should also induce the teams to invest in the training and development of players. As said earlier, professional soccer teams are usually responsible, through academies and affiliated teams, of the training and development of young players. In addition, they have to take care of the development of the players that are currently on their roster. The transfer fee is seen as the mean of getting back the costs of training the player that was transferred and the costs of training of the other players that were trained but did not become professionals. According to the traditional soccer economic theory, the transfer fees that a team receives are a way of reimbursing these expenses linked to the development system.

Since this theory was developed to describe the pre-\textit{Bosman} situation, one must conclude that the ECJ did not completely agree with the traditional sports’ economic reasoning, at least in the case of the players at the end of their contracts. It judged that the competitive balance was not helped with the existence of transfer fees, based on evidence presented by Bosman’s lawyer that showed that the rules did not prevent the richest teams from getting the best players or help the poorer teams get better on-field performances. These facts have also been confirmed by other sources.

As for the second argument, the reimbursement of the training and development costs, the ECJ found, again using evidence presented by Bosman’s lawyer, that the transfer fees were contingent, uncertain and totally unrelated to the actual costs of training a player. Consequently, the prospect of receiving such fees could not be an incentive for any team to invest in the development of players.

Finally, it found that the same aims could be achieved at least as efficiently by less restrictive methods which impeded the free movement of persons less. The ECJ with this short commentary rebuked the traditional economic explanation of the transfer fees proposed by the sports’ economists. But it recognized both goals as legitimate by declaring that any measure that achieved them would have to be considered immune from the effects of European law. There is still a debate if both these goals must be met or if only one would be enough. I am of the opinion that both goals must be met since immunity from EU law is an important benefit.

\textsuperscript{259} Bosman, paragraph 107, see note 9.  
\textsuperscript{260} MORROW, S., 1999. see note 50.  
\textsuperscript{261} Bosman, paragraphs 107 to 110, see note 9.}
that should not be granted lightly. But it is possible that meeting only one of these two conditions could be enough. Further decisions from the ECJ will be needed in order to settle this debate.

5.9.4. The Monti rules

In 2001 FIFA finally introduced the Monti rules as the new transfer system applicable to the world of soccer.

But in 2005, the FIFA executive committee, after hearing numerous complaints from teams and national associations regarding the complexity of the Monti rules, adopted a new, simplified system that is in effect today. In the following sections I will first examine and analyze the Monti rules. Although no longer in effect, they are still reflected in the actual rules. It will allow the reader to better understand the legislative process of FIFA and the legal reasoning that lies behind the rules, focusing on the differences of the actual rules with the Monti ones in order to better understand their effect on the value of players.

The rules cover both amateur and non-amateur players, but for the purpose of this analysis I will only examine the non-amateur transfer rules since they are the ones affecting professional players. These rules consist mainly of two documents: the Regulation for the status and transfer of players (Transfer Regulations) and the Application regulations.

First, the rules set up a limit to the length of professional players’ contracts. According to paragraph 4 (2) of the Transfer Regulations, contracts should have a term of between one and five years. This seems to have been adopted to counter one measure taken by teams following the Bosman decision: signing up their players for very long periods of time. This provision was put in place to prevent teams from binding their players for an unlimited or unreasonably long period of time to prevent them from ever being free agents.

Then the rules go on to regulate the transfer fees per se. Articles 13 to 20 of the Transfer Regulations contain the provisions that regulate the compensation for the training and


263 Regulations for the Status and Transfer of Players (Transfer Regulations), FIFA Executive Committee, July 5 2001 and Regulations governing the Application of the Regulations for the Status and Transfer of Player (Application Regulations), FIFA Executive Committee, July 5 2001.
education of young players. The aim of these provisions is to make sure that transfer fees are related to the costs of training of a young player and that they are paid to the teams that trained them. Also, it designates the years between the age of twelve and twenty-one as the years in which the education of a young player occurs and that a compensation for the costs of training the player should be paid in all cases if he is transferred during that time. But there is an inconsistency with that goal in article 13. The article states that if it is evident that a player has already ended his training period before the age of twenty-one, compensation shall be due until the player reaches the age of twenty-three, but that the fee shall be based on the years between twelve and the age at which it is determined that the player ended his training period. So instead of having a hard set rule, we have a soft one that leaves a lot of discretion to the teams since they will most likely be the ones that determine if and when a player ended his training period. Teams could be tempted to use that article in order to capture most of the future transfer fees attached to this player for themselves. By unilaterally declaring the training period of a player to be over when he leaves the team, that would crystallize the value of the training and development period to make sure that they will not have to share the fees for transferring him with other teams. Finally, article 17 outlaws any other type of compensation or transfer fee for players under twenty-four except for the amount prescribed by the Transfer Regulations.

The formulation of the Transfer Regulations is very general and we need to look at the Application Regulations in order to try to completely understand the process. However, these additional regulations still do not answer all our questions.

In Article 6 of the Application Regulations one can find the method with which the compensation is calculated.

First, it establishes four categories that determine the compensation to be paid for a player. These are

Category 1- division 1 teams of major soccer countries
Category 2- division 2 teams of major soccer countries and division 1 teams of other countries
Category 3- division 3 teams of major soccer countries and division 2 of other countries
Category 4- all the other teams
The determination of the team category shall be made by the National Associations of the various countries member of FIFA in partnership with the players’ representatives. The National Associations are granted a lot of power in the determination of the compensation for training and education. Each year, for each category, except category 1 where it shall be based on the real costs of training a player, the National Associations shall determine a ceiling for the training costs of a player. So to determine the transfer fee, the costs of training in that category of team will be multiplied by the number of years the player stayed with the team of that category. Then all the years of training in all the teams the player has played for are added together to give the total compensation payable for the transfer of the player.

Article 8 of the Application Regulations determines how the compensation is distributed between the teams that contributed to the training and education of the player. If the transfer is made from a team of the category 3 or 4 to a team of a higher category, 75% of the amount shall be distributed to the teams that trained the player, pro-rated by the years passed in each such team. It diminishes to 50% if the transfer is from a Category 2 team to a team of Category 1 and to ten percent in the case of a player transferred between teams of the same category. The rest of the transfer fee goes to the team that transferred or “sold” the player, in addition to what it can claim as compensation for the training of the player. This results in some quite complicated bookkeeping for each transfer.

In addition to that, chapter IX of the Transfer Rules creates a solidarity mechanism that is added to the transfer system: according to that mechanism five percent of the total fee paid in all transfers will be paid to all the other teams that trained the player, in a proportion determined by the age of the player at the time he was with that team. So there is a double compensation for the teams that trained the player, first through the amount that is paid directly to them according to Article 8 Application Regulations and then from the amount received from the former team through this “solidarity tax”. This solidarity tax is not related to the age of the player and applies to all transfers payments, not just the training compensation payment.

Finally, it is interesting to note that the transfer fees are not declared illegal for the players that are above twenty-three years of age. These players are still subject to the post-Bosman rules and they are likely to be transferred based on their economic market value and based on

---

264 Application Regulations, Article 10.
the cost of training them. If the players are without a valid contract of employment, they can freely negotiate their services to all teams.

5.9.4.1. Analysis of the Monti rules

It is clear, reading the Monti rules, that they were specifically designated to prevent another Bosman case. However, it is very surprising to see the number of holes that were still present in the regulations and that could have resulted in legal contestation.

The Monti system tries to give more incentives to the teams to train and educate new players. There are incentives on two fronts, first the money received each time a player is transferred and secondly the cost of the player to a team that decides to acquire him instead of trying to “grow” new talent. Nevertheless, this system does not guarantee additional investment in training and development.

But the main problem of the Monti rules is that the determination of the costs of a year of training has to be done by the National Association of each country. This creates a monstrous bureaucracy. By letting the National Associations and not the market decide the costs of the compensation for training and education, FIFA puts the burden on them to find the efficient level of compensation that will allow the optimal level of training. It is a difficult task and it would seem a better idea to let the market decide the value of a year of training in a team of a particular category or division in a particular country. Negotiation between the teams would make it possible to determine the market value of that year of training. Two teams could be in the same category but the expertise of one could make a year in that team a lot more valuable. Some kind of supervision would be needed from FIFA and the National Associations to make sure that there is no abuse in the system and to provide a conflict resolution system. Nevertheless, by leaving all the weight of the decision on the National Associations, the Monti rules run the risk of under-valuating the costs associated with the training of the players.

It is also very unfortunate that the Monti rules, in Article 5(5) of the Application Regulations, still require compensation for a young player without a contract. This seems to be in

---

265 Application Regulations, paragraph 7(5).
266 FEESS AND MÜHLHEÜBER, see note 75 and ANTONIONI AND CUBBIN, see note 253.
contradiction with the *Bosman* ruling. Although the ECJ said that some kind of restrictive behaviour may be acceptable to incite teams to invest in the training of players, the vagueness of the AR that ask to “take into consideration” the absence of contract, but without prejudice to the other teams that were implicated in the training of the player, is very dangerous. It is hard to know what the impact of this provision will be since there are many ways of taking something “into consideration”. It is difficult to believe that it was impossible to create a precise criterion, like a discount percentage, that would allow a clear and objective evaluation of the absence of contract for a young player’s transfer fee. A discount of 25% to 50% of the transfer fee, depending on the age of the player and the category of the teams implicated in the transfer, could have been a solution to make a player more attractive for the other teams and prevent another *Bosman*-like situation. This would have the advantage of making the situation clear for all those involved in the transfer process.

Regarding the second part of the test set up by the ECJ on the validity of the rules that restrict the market for players, I do not see how the Monti rules will help to maintain the competitiveness of the game in the commercial private teams market. Rich teams will always be able to pay the compensation for training in order to get the players that they want assuming that the gap in revenues between rich and poor teams remains constant and revenues continue to rise at their current rates.\(^{267}\) However, even if the commercial private teams’ competitive balance remains unaffected, the greater freedom for the players could benefit the national soccer teams. It has been\(^{268}\) argued that the quality of international soccer has improved since players are allowed to play in better foreign leagues. Again, in deciding which competitive balance to promote, FIFA must choose the type of soccer that it wants to encourage. But it is uncertain whether the ECJ would accept that view since in *Bosman* it mainly looked at the competitive balance at the team level and not at the national teams level. But the argument of approaching competitive balance from another perspective is not without merits and should be at least considered.

Finally, some may find peace in the fact that the regulations were approved by the Commission, but this is no guarantee of validity since the rules that were invalidated by *Bosman* had also been approved by the Commission.\(^{269}\)

\(^{267}\) DOBSON, S. AND GODDARD, J., see note 47.


\(^{269}\) *Bosman*, paragraphs 17 to 27, see note 9.
5.9.5. The 2005 transfer rules

Almost from the moment of their implementation, the Monti rules were severely criticized by the various soccer national associations and professional teams. Their main complaint was that it forced them to create additional levels of bureaucracy to manage the compensation system and that it required complex bookkeeping. They had the impression that the solution was worse than the original problem. Faced with this resistance, FIFA decided to modify the Monti rules. The resulting adjustments gave us the 2005 transfer rules, a simpler leaner system that tried to retain the main characteristics of its predecessor and still comply with Bosman.

5.9.5.1. The changes

Regarding the legal length of a player’s contract, the 2005 rules have kept the same duration as the Monti rules, but they have added a provision regarding contracts with minors: they are now limited to three years in duration. Again, the aim seems to be to prevent long contracts with minors intended to secure transfer fees by all means possible by the teams.

The transfer fees are now called, for the purpose of the regulation, training compensation. There are two cases that allow a team to get training compensation for a player. The first case is when the player signs his first professional contract and the second one when the player is transferred between teams in two different national associations which could also be called an international transfer.

When the player signs his first professional contract, training compensation will be paid to all the teams that trained him between the age of twelve and twenty-one, at the pro-rata of the years he passed in the said team. That covers all teams that trained him as an amateur.

In the second case, the international transfer of a player that is already a professional, the team that acquires the player will pay training compensation only to the former team for the

271 Interview with Pedro Trangrouse, Vice-President Legal affairs of the Brazilian Football Federation, November 2005 and interview with Alan Vessey of the Football Federation Australia, see note 56.
272 2005 Transfer rules, paragraph 18 (2).
273 2005 Transfer Rules, Annex 4, section 2 (1).
duration of time that it trained the player.\textsuperscript{274} There is no complex kickback scheme like in the Monti rules but this is a straight forward one shot payment. It is important to notice that the 2005 rules do not exclude additional payments on top of the training compensation when a player under the age of twenty-four is transferred. The prohibition on payment from other compensation of article 17 of the Monti rules was not reproduced which gives teams more freedom in determining the costs of a player.

For the training compensation in the case of transfer within an association, the national associations were left in charge of determining it, subject to the approval of FIFA.\textsuperscript{275} The associations had until 2007 to adopt these rules and submit them to FIFA. The only guideline that can be found in the 2005 rules is that the national rules must be compatible with the ones of FIFA.\textsuperscript{276} The actual situation in the different national transfer markets is relatively confusing; some associations missed the deadline whereas others like the Australian Federation decided to make every national transfer free and still others decided to simply copy the FIFA system at the national level.\textsuperscript{277} The 2005 rules abolished training compensation when a transfer happens between two teams of category four.\textsuperscript{278}

In addition, if it is impossible to know or to find one or more teams that trained the player, then the money due to that team will be given to the National Association of the country in which the player was trained. This money should be earmarked by the association for the training and development of young players.\textsuperscript{279}

Finally, the classification of the teams in the different categories and the determination of the cost of training for each of them are now determined by FIFA on a confederation wide basis\textsuperscript{280} and should be based on the real training costs of training a player, multiplied by the number of players needed to produce one professional player. As for the categories, FIFA publishes the number of categories that the national associations need to divide their professional teams into.\textsuperscript{281} National associations may have only a restricted number of

\textsuperscript{274} 2005 Transfer Rules, Annex 4, paragraph 3 (1).
\textsuperscript{275} 2005 Transfer Rules, paragraph 1 (2).
\textsuperscript{276} 2005 Transfer Rules, paragraph 26 (3).
\textsuperscript{277} Interview with Pedro Trangrouse, see note 271 and Alan Vessey of the Football Federation Australia, see note 56.
\textsuperscript{278} 2005 Transfer Rules, section 2 (2) ii).
\textsuperscript{279} 2005 Transfer Rules, paragraph 3 (3).
\textsuperscript{280} Annex to FIFA Circular 959, FIFA Circular 1085 and FIFA Circular 1142.
\textsuperscript{281} 2005 Transfer Rules, paragraph 4 (1).
categories like Canada with two (the third and the forth) or have access to all four like Germany, Brazil and England\textsuperscript{282}. The training compensation is then simply determined by the amount of compensation due for one year according to the category of the new team and then multiplied by the number of years the player has been trained in the old team.

The great change of the 2005 rules is that the five percent solidarity tax will now be charged in all cases where a non-training compensation fee is paid for the player at any time during his career. The term non-training compensation fee is not defined in the rules but we may deduce that it means any fee paid for the transfer of a player that is determined by the training compensation regulation. This confirms that the transfer fees for the economic value of the player are not illegal under the 2005 rules. They are now even legal in the case of the transfer of a player covered by training compensation since article 17 of the Monti rules that prohibited them was not reproduced in the 2005 rules. The above mentioned tax will apply to all international transfers and then be kicked back, according to a defined scale, to all teams that participated in the training of the player.\textsuperscript{283} This kickback system is simplified to the point that it is a simple pro-rata of the number of years that the team trained the player. It does not compare to the complicated system in the Monti rules and is rather straightforward and therefore easy to understand.

5.9.5.2. Special changes for Europe

Annex 4, Article 6 of the 2005 transfer rules creates special provisions for transfers done in the EU.

Paragraph 6 (1) states that when a player is transferred to a team of a higher category, the training compensation will be calculated using the average between the costs of both teams. If the player is transferred to a lower category, then the training compensation will be the one of the new team in the lower category.

Paragraph 6 (2) allows a team to determine the end of the training period of a player before the mandatory age. It is surprising to see that one of the main legal problems of the Monti rules survived the review, especially in the market where a judicial contestation is most likely

\textsuperscript{282} Annex to FIFA Circular 959, FIFA Circular 1085 and FIFA Circular 1142.
\textsuperscript{283} 2005 Transfer Rules, Annex 5, Article 1.
to happen. The purpose that this paragraph will serve remains unclear; as stated earlier in the analysis of the Monti rules, it seems to have been designed to allow teams to manipulate training compensation for their own purpose.

But on the positive side, Paragraph 6 (3) seems to solve the problem of the restriction of movement of players under twenty-four without a valid contract. If a player is without a valid contract, no training compensation is due to his last team unless it can justify that it is entitled to such compensation. However, this does not affect the rights of the other teams that trained him in the past; they retain the right to training compensation for the years the player was a member of their team. Although not free of future potential legal problems, this change is at least a step in the right direction compared to the complete restriction that was found in the Monti rules. Any restraints on players without a contract can create problems, but this solution is less problematic than the one contained in the Monti rules. Teams that trained the player should not be penalized because of a conflict between the player and his current team; their rights to compensation should not be lost depending on how much this would in fact restrict the movement of the player. Therefore this restriction could be justified.

So after all these legal rewordings and adjustments, what impact do the new transfer rules have on the value of players?

The 2005 transfer rules introduce a simplified method of determining the training compensation. It is refreshing to see that at least now we have an objective table to determine the costs of training in the FIFA Circulars 1142, 1085 and before that in the annex of Circular 959\(^\text{284}\). But it is still unclear if these amounts could be considered objectively calculated and related to the costs of training in case of contestation. Do they give teams an incentive to correctly value the players and pay them accordingly? Finally, do they affect the competitive balance in soccer?

5.10. Effect of the transfer rules on the value of soccer players

Optimal investment in the training and education of the players is in the interest of all parties implicated in the soccer business. It allows teams to replace the players who must retire and raises the level of competition since there is a limited number of players that a team is

\(^{284}\) All available on the FIFA website at http://www.fifa.com.
allowed to field. If the number of players fielded stays the same but the number of “fieldable” players grows, there will be a better quality of players in all the teams, making the matches more exciting. This should generate more demand for the professional soccer product and consequently more revenues for all teams, or at least for the industry as a whole. Assuming that the percentage of revenues given to players will stay constant, it is also in the interest of players that will see their aggregate value grow.

But in order for the new transfer rules to be effective, the compensation received by the teams for a player must be sufficiently high and occurs frequently enough to:

1- compensate the actual costs of training the player
2- compensate the costs of training the proportionate number of other players that will not reach the professional level to the number of players that do
3- be proportionally and sufficiently higher than the positive externality of training to third party teams coming from the investment made in training by the team

If the new rules achieve these conditions, they will create an incentive for the teams to invest in training. If the fees are at the efficient level, all teams that want to maximize their amount of player talent will be induced to invest in the training of players in order to “grow” some of the talent that they need, sell excess talent and buy the players that they were not able to develop. By allowing transfers we allow the allocation of the players from the teams that are the most efficient to train them to the teams that value them the most. This is to the benefit of players that are allowed to move to teams that provide them with the better pay or playing opportunities. The incentive game is subtle and the amount of the transfer fees should be decided with great care.

Looking at the actual tables contained in the FIFA circulars, one feels that something is wrong. The 2007 and 2008 tables are exactly the same as the one published for 2005, although it is clearly written in all circulars

“...In accordance to Article 6 of the Regulations for the Status and Transfer of Players, these training costs will be updated at the end of every calendar year.”

---

285 Assuming an equitable revenue distribution.
286 Annex to FIFA Circular 959, FIFA Circular 1085 and FIFA Circular 1142.
Does that mean that the training costs in soccer are immune to inflation? But as one looks at the numbers in the circular, it gets even worse: although costs are supposed to be determined for each confederation, there are very little differences between them as the table below shows.

Table 2

<table>
<thead>
<tr>
<th>Confederation</th>
<th>Category I</th>
<th>Category II</th>
<th>Category III</th>
<th>Category IV</th>
</tr>
</thead>
<tbody>
<tr>
<td>UEFA (Europe)</td>
<td>€ 90000</td>
<td>€ 60000</td>
<td>€ 30000</td>
<td>€ 10000</td>
</tr>
<tr>
<td>CONMEBOL (South America)</td>
<td>$ 50000</td>
<td>$ 30000</td>
<td>$ 10000</td>
<td>$ 2000</td>
</tr>
<tr>
<td>OFC (Oceania)</td>
<td>$ 30000</td>
<td>$ 10000</td>
<td>$ 2000</td>
<td></td>
</tr>
<tr>
<td>CAF (Africa)</td>
<td>$ 30000</td>
<td>$ 10000</td>
<td>$ 2000</td>
<td></td>
</tr>
<tr>
<td>CONCACAF (North America + Caribbean)</td>
<td>$ 40000</td>
<td>$ 10000</td>
<td>$ 2000</td>
<td></td>
</tr>
<tr>
<td>AFC (Asia)</td>
<td>$ 40000</td>
<td>$ 10000</td>
<td>$ 2000</td>
<td></td>
</tr>
</tbody>
</table>

(Amount given in Euro or US Dollars)

According to that table, the costs of training and developing young players are pretty much the same across the world, except for Europe, where it is between 55% and 500% higher, before factoring in the difference in value of the currencies. There is another problem with the circular: the method that has been chosen by FIFA to determine the costs of training does not consider the variety of countries included in the various associations. Taking the example of UEFA, is it realistic to consider that the training costs in Germany, Russia and Albania are similar? The same thing may be said about the countries of the CONCACAF which include the United States, Central America and the Caribbean. One has the distinct impression that there was no serious research done by FIFA to determine objectively the real costs of training.

287 Here is, for the benefit of the reader, the list of the 53 members of UEFA: Albania, Andorra, Armenia, Austria, Azerbaijan, Belarus, Belgium, Bosnia and Herzegovina, Bulgaria, Croatia, Cyprus, Czech Republic, Denmark, England, Estonia, Faroe Islands, Finland, France, Georgia, Germany, Greece, Hungary, Iceland, Republic of Ireland, Israel, Italy, Kazakhstan, Latvia, Liechtenstein, Lithuania, Luxembourg, FYR Macedonia, Malta, Moldova, Montenegro, Netherlands, Northern Ireland, Norway, Poland, Portugal, Romania, Russia, San Marino, Scotland, Serbia, Slovakia, Slovenia, Spain, Sweden, Switzerland, Turkey, Ukraine and Wales.

288 Again, for the benefit of the reader, the members of CONCACAF are: Anguilla, Antigua and Barbuda, Aruba, Bahamas, Barbados, Belize, Bermuda, British Virgin Islands, Canada, Cayman Islands, Costa Rica, Cuba, Dominica, Dominican Republic, El Salvador, Grenada, Guatemala, Guyana, Haiti, Honduras, Jamaica, Mexico, Montserrat, Netherlands Antilles, Nicaragua, Panama, Puerto Rico, Saint Kitts and Nevis, Saint Lucia, Saint Vincent and the Grenadines, Suriname, Trinidad and Tobago, Turks and Caicos Islands, USA and U.S. Virgin Islands.
and developing young players. If we compare these costs with the amounts invested privately by teams in the major North American sports, it is hard to believe that they are realistic. National Hockey League teams are said to spend between one and three million dollars (US) each year to support their development system\textsuperscript{289}. That gives us around 50 000 and 150 000 USD per year per player, excluding the salaries paid to them. For Major League Baseball, an audit conducted in 1991 by Ernst & Young evaluated that Major League Baseball teams spent an average of 7, 2 million dollars each year on scouting and development costs\textsuperscript{290}. These examples from other sports make it hard to believe that FIFA’s numbers are realistic or have been the object of serious research on the real costs of developing players. But this is only a preliminary conclusion and more research is needed on that question to come to a definitive conclusion.

More empirical studies will also be needed to determine if the amount of money offered by the current system is enough to give the right incentives to train and develop young players. But at first glance, there seems to be a risk of under- and overvaluation of the costs of a year of training due to the confederation wide training costs system. The member countries in each confederation have too many differences in their socioeconomic situation for that “one size fits all system method” to be efficient. It is even worse than letting the national associations determine the compensation like in the Monti rules. At least, the national associations were closer to the real costs of training in their country than FIFA. With the current system, there are great risks that we end up with a situation in which there will be an imbalance in the market for training of players and teams, depending on the country that they are located in and the incentives that it offers to train players. Teams will have to choose if they become either selling or buying teams. This situation is unlikely to be efficient since this will result in a redistribution of good players from countries that developed them to the ones that buy them. It will however do nothing to help get the efficient amount of players developed since teams in countries where real costs of training are under the FIFA determined level will concentrate on developing players to sell them to countries where the real costs are above the FIFA determined level. Teams that are located in the latter countries will therefore stop developing their own players and buy them from low cost countries at a lower “turnkey” cost. Since the optimal levels of training and development and transfer are never 0 or 100%, the current situation therefore will result in an inefficient solution.

\textsuperscript{289} Interview with Julien Brisebois, see note 114 and Yves Vonlanthen, Administration & Legal Manager of the International Ice Hockey Federation, March 2007.

\textsuperscript{290} KRAUTMANN, A. AND OPPENHEIMER, M., see note 233.
A possible solution to correct this imbalance would be to let the market decide the individual value of training in each club under some sort of supervision from FIFA. There is a dispute resolution system included in the 2005 rules that has jurisdiction to adjudicate on training compensation should the implicated teams not agree on it. Since the international transfers need to be reported to FIFA, it would have been easy to create a monitoring system to prevent abuses and help create a training compensation scale that would give the right incentives to teams to properly value the players.

In addition, it is a relief to see that the complete restriction on the movement of young players without contracts contained in 5 (5) AR of the Monti rules was modified. This visible contradiction of the Bosman decision is now replaced by a less restrictive system. According to paragraph 6 (3) of the 2005 rules, the current team does not have the right to a compensation for training if the player does not have a valid contract. If the team wants compensation, it now has the burden to prove that it nevertheless has the right to it. But there are no real criteria in the 2005 rules that help determine what could give the right to a team to ask for compensation in the case of a player without a valid contract. Will teams be forced to prove the real costs of training the player? With time, we will hopefully be able to see some criteria come from the dispute resolution system, but there are still none available in the Dispute Resolution Chamber (DRC) jurisprudence. Our only guide is the following statement in the commented regulations, a document published by FIFA in 2007, which states

"3. If the former team does not offer a professional player a new employment contract, this team loses its entitlement to training compensation unless it can justify that it is entitled to such compensation. This justification may be very difficult to prove and limited to extraordinary circumstances to decide on the matter at hand.

4. In order to safeguard its entitlement to training compensation and demonstrate its real intention to continue its relationship with the player concerned, the former team must offer the player a contract in writing via registered mail at least 60 days before the expiry of his current contract. The offer in the new contract shall at least be of an equivalent value to the current contract, otherwise it is as if the team did not offer a contract at all,

---

293 Commentary on the Regulations for the Status and Transfer of Players, FIFA Executive Committee.
with the consequence that if the player moves to another team within the EU/EEA, no training compensation is payable to the former team.  

At least, this comment clears two things: first, the application of that provision is intended for exceptional cases only, and secondly, if a team wants to use it, it must have made a new contract offer to the targeted player that is at least equal to the value of the one that just expired. But one must note that giving a player a contract equal to his last one does not mean that the player is going to be satisfied with that offer. So the main problem that led to Bosman is still potentially present in the new rules, especially for young players that reach the peak of their career and performance.

Another interesting point is hidden in a footnote of the Commentary concerning that article within which FIFA gives us an example of a situation when a team could receive a fee without offering a contract to the player

“If a team descends to a lower division in which it is not entitled to register players as professionals, this team will not be in a position to offer an employment contract to young players. However, it will not forfeit its entitlement to claim for training compensation from the player’s new team.”

This is very interesting and indeed looks like a situation in which it is reasonable to allow a team to receive training compensation even if the player does not have a valid contract. It would probably meet the condition of the Bosman test. It is however disappointing to see that FIFA can only provide vague observations or single examples in these commentaries when the implementation of updated rules in 2005 would have been the perfect occasion to formulate some concrete legal guidelines or criteria to help teams and players to predict in advance the effect of this rule. In the current situation, too much uncertainty remains and the parties must wait for jurisprudence from the DRC in order to measure the true impact of this article. It is an unsatisfying regulation to leave the final word with the courts.

5.10.1. The direct influence of the new transfer system on the value of the players

By determining directly in its circulars how much a year of training is worth, FIFA in fact created a minimum floor value for every player in every confederation. This is a direct

---

294 Idem, p.125.
295 Idem, p.125.
influence on the value of players since, if we again refer to the equation in chapter 3, FIFA determines a minimum AC for each player. As there is no limitation or ceiling on the value of the training compensation, a rational team will always ask for at least the minimum value that is written in the FIFA circular. Any amount above that will be entirely discretionary and the result of the negotiation process between the teams. It is safe to assume that it will be based on the willingness to pay of the buying club. So in the case of a player where the willingness to pay of the buying club is greater than the minimum FIFA determined problem, the current system will not influence the value of the player strongly. In the end, the amount will be the same as if it had been determined by the free market. The problems start with players that have a lower market value than the FIFA determined minimum one. For them, any movement will be difficult unless the teams are ready to breach the regulations since otherwise, any transaction based on the FIFA value would result in a loss for the team acquiring him (ES will be negative). Therefore, the new FIFA transfer system has a direct influence on the individual value of marginal and average players by putting a floor to their value and hence, it can make their acquisition a loss instead of a benefit. The logical consequence, assuming that the teams respect the rules, will be an elimination of the international transfer market for these players.

The FIFA rules also have a redistributive effect since, with the establishment of the floor, they assign part of the value of AC based on the confederation the player was trained in. This forced assignment may not reflect the conditions of the market or the will of the parties, but as it is mandatory, it cannot be avoided and the parties must suffer its consequences. The logical consequence for the club forced to pay a higher transfer fee (AC) than the one he is willing to pay is to offer a lower salary (S) to the player. We therefore have a situation where FIFA created a redistribution of the value of the player in favour of the teams by allowing the selling team to capture some of the value that should have been, in a free market, attributed to the player. So the rules collectively redistribute more of the value of the sport to the teams and in consequence lower the aggregate value of the players.

5.10.2. The indirect influence of the rules on the value of players via competitive balance

This section is necessary for, as mentioned earlier, competitive balance also has a direct effect on the value of players. Therefore, if a commercial rule of the game has an influence on competitive balance, it will also result in a change in the aggregate value of the players.
looking at the direct effects of the new FIFA transfer rules, the competitive balance approach looks at their indirect effect.

It is hard to determine if the 2005 rules have a significant impact on the competitive balance of soccer. Like the Monti rules, they do not affect the main reason for imbalance in European soccer which is the huge gap in revenues between the different teams\(^{297}\). The five percent solidarity tax may be a step in the right direction, but it is only a small amount and will be distributed to all teams, irrespective of their revenues. It is not a way of subsidizing poor teams but a way of rewarding success in developing players. It may in fact aggravate the problem since richer teams with bigger development budgets can develop more players and get more solidarity tax. Affecting the competitive balance of a league is always difficult unless the problems at the basis of the imbalance are directly addressed. The solidarity tax may have an impact but only time will tell if it helps solve the imbalance.

Since the transfer rules do not address the main cause of imbalance in soccer, that is the great difference in revenues between the different teams\(^{298}\), they are unlikely to significantly affect the competitive balance of soccer. In addition, as they are only applicable in case of international transfers, they fail to influence the other important part of the transfer market, the domestic transfer.

If one assumes that the rules do have a positive influence on the competitive balance of soccer, then this would result in a greater aggregate value of the players, assuming that the percentage of the total revenues of the sport invested in salaries would remain the same. However, since they affect only part of the transfer market, such an influence is unlikely and yet if it is present, it will be only very small. I am convinced that more research is needed on that field in order to be able to clearly distinguish the indirect influence of the new rules on the value of players via competitive balance.

\(^{297}\) Szymanski, S. AND Zimbalist, A., see note 57.

\(^{298}\) Idem.
5.10.3. The instability of the new system

The new FIFA transfer rules are both full of constraints and opportunities for players. It is the first step in the right direction of recognising them as more than chattel and empowering them to take the opportunities offered in the modern soccer market.

But contrary to other employees, soccer players do not face a free market for their services. They are still forced by being part of a transfer system in which they can be sold, for a fee, to another team. This remains a unique situation since normal employees do not simply get “transferred” if another employer wants to acquire their services. Also, when young promising employees change jobs at the end of their employment contracts, their new employer does not have to pay the old one for the training and formation they received when working for him. For these reasons, the transfer fee system is one of the most important influences on the value of soccer players and it determines the main conditions of the market for their services.

Transfer fees rarely reflect the investment that a team made in the training and development of a player. As stated by the ECJ in Bosman, they are discretionary and the result of the negotiations between the teams involved in the transfer. The transfer fee reflects part of the market value of the player, as the equation proposed in chapter 3 shows. It represents the value the buying team places on the playing rights of the player. This value is assigned unilaterally by the teams to reflect the on-field and off-field commercial value of the player.

In consequence, with the new transfer rules, FIFA tries to control the market in the way it desires, for example by reflecting the cost of training in the transfer fee. So until the player turns 24 years of age, the rules will make him lose some of the benefit of his value by assigning it to the teams that trained him. By forcing the addition of the training and development costs into the equation, the new FIFA system will mainly result in a redistribution of the value of the player to the teams in order to meet these payments. The money will either come from the salary of the player or from the “market value” transfer fee as the new system is unlikely to increase the willingness to pay of the teams. Since the value of a year of training is fixed by FIFA, it could force a team to pay more than they it thinks the player is worth just to respect the rules. The logical consequence will be a lower salary for the player for there will be less money left in the “pot” for him. This will result in lower salaries
for the players, the money for the salaries coming from the same pool as for the other fees. FIFA assumes that this redistribution of money will give teams the incentives to invest in the training and development of players. More money to develop young players should result, in the long run, in soccer of better quality and in greater competitive balance for the sport. This should provide more revenues and, assuming that the percentages of it given to players stay the same, in an increase of the aggregate value of players. In short, the motives are the same as for a change in the rules of the game, at least in the case of the FIFA transfer rules. In theory all that may be true, but the structure of the current system makes these ideas hard to believe since it has numerous flaws.

Even if FIFA wanted us to believe it, it is hard to imagine that there is a link between a confederation wide determined amount and the real costs of training and development of players in individual countries of this confederation. As an example, the average salary in Romania is around seven times less than in Germany\textsuperscript{299} and both are members of the same confederation (UEFA). Teams from the rich countries in the confederation could find it much cheaper to buy talent from the poor countries than to actually invest in the training and development of their own young players. For the teams from poorer countries, there is an over incentive to develop young players in order to sell them abroad since a player is much more valuable if sold to a rich team than as an on-field performer. This is reinforced by the solidarity tax which makes it even more attractive to develop players since the former club will be paid a dividend every time these players are transferred in the future. Not giving the correct incentives to develop the efficient amount of players in all countries of the confederation, the FIFA rules do not result in an efficient market for the players. The teams of poor countries will have too much incentive to develop players and teams from rich countries too little. This gives us inefficient markets where there is over- and underinvestment in training. This is not just a theoretical problem for the players since this under and over investment will affect the way the player will reach his potential and the types of players that are developed.

For the players, the consequences of the rules will depend on the country they have been trained in. Depending on where they are from, the value attributed to them will be different. The goal of the teams in countries where the costs of training are cheap and that have an over

incentive to sell players will be to develop and sell the players before they reach the age of 24 years. They will try to find young players who will quickly reach their peak and not players that require patience and a long investment in developing them for if they take too long to develop, the team could lose them without compensation when they reach their peak at an age older than 24 years. This will create a redistribution of value between the different types of players with the young and quick to develop players being more valuable than the ones needing more time to mature.

For the teams with real costs of training higher than the FIFA determined ones, their incentive will be to buy players. They will not have any incentive to sell any of the players that they develop themselves for the fee that they would receive would not even cover the actual costs of training them.

The same imbalance is reflected with the global system of compensation determined by FIFA for international transfers: it gives too much incentive for the teams outside Europe to train and develop young players with the objective of transferring them to Europe. There is however a limited supply of players talented enough to play in the European leagues. This leads to an over investment from teams in recruitment to find these players and to make sure that they reach their full potential before they reach 24 years of age. So these Non-EU countries have to be added to the poor European countries that also want to sell their players in the same market. With too many teams chasing a limited supply of players, the investment turns out to be higher than the efficient amount that should be spent and more than the efficient amount of money will be spent by the teams in scouting, training and developing players in the hope of finding the players that will provide high revenues.

In addition, teams outside of Europe have little incentive to transfer their players inside their own countries since the national transfer rules usually do not offer such an interesting return on their investments. The countries that are exporters make good business selling their players to the rich European teams. So for them, the new rules effectively create two values for their players - their value inside the country and their value in Europe. Both values are artificially determined by either the FIFA or the National Association and not linked with either the market value of the player or the real costs of the team to train and develop them. Players that could make attractive careers elsewhere or in their home country are not that

---

300 Interview with Pedro Trangrouse, see note 271 and Alan Vessey of the Australian federation, see note 56.
interesting to develop and their value is dragged down by the transfer system that puts the
emphasis on the development of players sold to European teams.

So these new rules have changed the way that players are valued, but not necessarily in a
manner that better reflects their playing capacities. The *Bosman* case and the changes in
transfer rules that followed it modified the way the monopolistic rent of soccer is
distributed. In the pre-*Bosman* era, due to the players’ inability to test the market for their
services, their salaries were kept artificially low. This was in accordance with the traditional
business model of professional sports that resulted in the players being kept out of the
distribution of the monopolistic surplus of the industry. The teams and the fans ended up
splitting most of it between themselves. For teams, it took the form of higher profits through a
combination of lower expenses and higher revenues. The low salaries of players resulted in
lower expenses and the transfer fees in higher revenues. For fans, it came in the form of low
ticket prices and the possibility of watching games for free on television. In the post-*Bosman*
world, the monopolistic rent now appears to be shared between the teams and the players with
the share of the players growing each year. The fans have lost their part of the rent since they
now pay more for the stadium tickets and also have to pay to watch games live on
television. The absence of transfer fees for players at the end of their contract means that
the money that was, under the old system, paid to the teams to acquire the players is now
available for their salaries. This is confirmed by the recent fast pace growth of players’
salaries after the *Bosman* case.

The reserve clause or, as later research has shown, any restriction on the salaries or free
movement of players, allows the teams to keep part of the rent of the players for themselves
by paying a lower salary. In the post-*Bosman* world, the players are the beneficiaries since
they can ask for higher salaries.

We can now see a redistribution of the monopolistic rent based on the new balance of power.
By having the possibility to offer their services to all teams in the world, the players are able
to play with the team that values them the most. Again using the equation of chapter 3, since
the acquisition cost (AC) of the player is lower, more money is available for his salary (S).

---

301 DOBSON, S. AND GODDARD, J., see note 47.
302 Idem.
303 Idem.
304 KÈSENNE, S., see note 20 and GRAVES, J., see note 231.
The implementation of the new transfer rules also results in a redistribution of value between the players. Players under 24 years of age now have a minimum value that could be interpreted as a minimum price or a ceiling price, depending on the costs of training in the country where they were trained. This value is given by FIFA independently of the playing ability of a player, the only criteria being the country and the category of the team that developed and trained him. As any price determined by a regulating authority, this is an inefficient solution since it is almost certain that the price determined by FIFA will not be similar to his price in a free market since; it forces a redistribution that may not be desirable and is based on criteria that have nothing to do with the real market value of the player\textsuperscript{305}.

In this specific case, the method of determining the price of a year of training used by FIFA compounds the problems by applying confederation wide prices to different countries with different supply, demand and prices. We therefore have a situation in which the FIFA determined price can and will never be close to the real costs of training and development in any country of the confederation. Such a determination will make the market value of the training and development of players in a country vary wildly, depending on the difference between the real costs of training a player and the costs determined by FIFA. Countries with real costs lower than the FIFA determined ones will have the incentives to export their players to countries where the real costs are higher than the FIFA determined ones. So players that have more potential to be sold will attract more attention and will have more value since they can be sold for a profit by their teams. This gets even worse when players are sold between federations. Let us imagine the following example: a team in South America (CONMEBOL) sells a good young player to a top European (UEFA) team, both teams of category I. By selling the player, the South American team will get 90 000 € per year of training which is around double the amount, before currency adjustment, the team would get if it sold him to a team of the same category in their own confederation (50 000 USD)\textsuperscript{306}. So the exportable player has two times more value if he is sold, not because the South American training is more efficient than in Europe but simply because the new FIFA rules say so.

\textsuperscript{305} PINDYCK, R.S. AND RUBINFELD, D.L., see note 13.
\textsuperscript{306} FIFA circulars, see note 286.
For the player himself, the European team, having paid a great amount of money to acquire him, will have less money available for his salary. So the FIFA rules redistribute the money from what the player should get to the team that trained them.

The inverse effect is also possible: a team with real costs of training higher than the FIFA determined ones will stop training all but the very best players who have more than an average value since it will lose money on training average players instead of simply buying them from other teams in countries with lower real costs.

For all the efforts FIFA put in the new system, it looks like we are facing a situation in which these restriction to the free movement of young players are not at all related to the real costs of training them.

What FIFA has done is try to partially restore the distribution of the monopolistic rent that existed before Bosman. But contrary to the example of the NFL earlier in this chapter, there was no negotiation with the players; FIFA imposed the rules that it considered appropriate. Only time will allow us to tell how the situation will evolve, but I am quite confident that it will also result mainly in a redistribution of the players’ part of the monopolistic rent between different categories of players. It will probably result in a very similar situation as in the NFL, with the important exception that the redistribution has been imposed by one party instead of having been negotiated by both. Only FIFA’s concerns are really reflected in the new rules. It will be interesting to see if the players are willing to tolerate that solution as they were ready to do with a negotiated CBA settlement in the case of the NFL.

If the players are not willing to tolerate the proposed solution, it is very probable that the transfer rules will again be contested in front of the courts. The result of such lawsuit is very hard to predict. However, as I have mentioned many times during my analysis, there are good arguments that can be provided against the legality of the new system. Imposing the new system made FIFA vulnerable to another ECJ decision that would force it to design yet another transfer system.
5.11. Conclusion: The impact of the change in the rules of the commercial game on the value of players

As we have seen, the effects of the commercial rules of the game on the value of players are essentially the same in two very different sports since the rules examined here were essentially aimed at the same goal: to restrict the mobility of players. Such rules have two main effects on the value of players.

First, the commercial rules of the game can divide the players into different classes. The best example is a free agency system that gives contractual freedom only to certain players. These classes operate a redistribution of value between the players themselves. Such classes will allow some players more freedom and in consequence allow them to negotiate better salaries for themselves at the expense of the ones that have restricted freedom of movement and negotiation. Like the rules of the game, the fact that a player has complete or restricted possibilities to offer his services to all the teams in the league will affect the value that he is able to secure for himself. But, unlike the rules of the game, he will be able to know in advance the effect the rules of the commercial game will have on him since they are already clear and determined when he start his career. He can plan for it, making sure that his contract ends at the best possible moment to maximize his negotiation position.

As said earlier in chapter 3, the main commercial operation of a sports league is to exploit the monopolistic rent coming from its dominant position in the market. One of the impacts of a change in the rules of the game is to change the way the rent is distributed between the different parties of the sport business - the players, the fans and the teams.

Restricting the ability of players to move between teams and their power to market their services gives them less than their share of the monopolistic rent than if it was a free market since it allows the teams to artificially lower the players’ salaries. The tendency that has been observed is that the greater the freedom of players, the more of the monopolistic rent of the sport will end up in their hands. Combined with the sportsman owner effect, this may result in the players getting all the rent, leaving the teams without profit and the fans having to pay very high amounts of money to enjoy and support their team. This may turn out to be an unsustainable business model for the sport business.
One of the solutions to that situation is to go the way the NFL went with a negotiated solution. This allows the teams and the players to determine themselves how the rent is going to be distributed between them. This creates an equilibrium in the redistribution that everybody agreed with, except for the fans that were not asked and end up having to pay more and more to see their teams perform.

In the case of the rules of the commercial game in the world of sport, either imposed or negotiated, the result is clearly a redistribution of the wealth of the business between the parties or between the different members of a party. But contrary to the rules of the game where the redistribution of wealth is a side effect of the imposition of these rules, the rules of the commercial game are usually directly aimed at affecting this redistribution. The point of adopting such rules is to shift the redistribution in one direction or another. The legislator of sport knows what he is doing when implementing the rules of the commercial game and making a business decision out of it. So we can conclude this chapter by saying that the point of the rules of the commercial game is to affect the redistribution of wealth and value between the different participants of the industry. For the players, any change in these rules has the potential of changing their individual and aggregate value.

If the rules of the game were a lottery for the players, then the rules of the commercial game would be the main instrument to determine what part of the monopolistic rent of sport they are allowed to appropriate. So this is why there is always much tension when rules of the commercial game are about to be changed since they have the potential of greatly affecting the revenues of each party concerned.

Depending on whether the rules are negotiated or imposed, the players have a certain degree of influence on the redistribution. In the case of negotiation, they are able to become partners with the teams in the business of sport since they get to determine how the revenues will be distributed, both within themselves and between the teams and the players. The rules of the commercial game allow, at least in certain cases, the players to determine part of their value and the size of the rent that will be attributed to them. With that, it is now time to draw the final conclusion of the different rules that I examined.
Conclusion

“We did not lose, we just ran out of time”
- Vince Lombardi, legendary NFL coach

In the recent years, the business of sport has started to embrace commercialism. This has been a long process since the end of the nineteenth century when any form of professionalism was strictly prohibited. But it was only a question of time before a team decided to pay the best players to play for it and that these players may come from a working class background that did not provide enough for them to play the game for free. The growing interest of the working class for organized sports provided the top teams with an audience that turned out to be willing to pay to look at them play. So with the advent of professionalism and paying public, the embrace of commercialism by the different parties in the business of sport was just the next logical step in order for teams to be able to compete on the commercial field like they do on the pitch. But even then, the business of sport refused to assume their pure commercial side. Teams and sport organizations have continued to claim that they have a special status and that they should not be treated as pure commercial venture and enterprises.

As we have seen, the claim of this special status is not uninterested. Parties in sport mostly want to use this status to operate outside of the normal laws and regulations. So is sport special after all? Let us shortly come back on the different issues that were examined in this work.

In the first chapter, we discovered that the structures of professional sport are comparable to a standard business model. First I showed that there may be some specificity to the business of sport, this specificity being linked to one of the product of the industry: the game. A game is a product in which one needs a competitor and where the standard optimization model of monopoly is the worst possible scenario for a team. Only looking at the business model of the single game and the revenues that are assigned to it one can claim: sport is special.

But the game only constitutes part of the equation. By analyzing further the model of the business of sport I showed that the market which is in fact exploited by the sport organizations is not the market for games but the market for championship. This market is usually exploited using the league as an organization. By looking at both the historical
evolution of the leagues and their economic nature and structure, I was able to show that in
the case of the exploitation of the market for championship, sport organizations have behaved
in a very standard way. In this case, the revenue optimization model is the same as for any
other organization: the monopoly. In addition, from a historical perspective, the first
professional leagues were set up by their founders in order to mimic the normal commercial
set up in the end of the XIXth century. Therefore, saying that sport is special is not accurate
from a commercial structure perspective. Certain elements of the industry may be specific,
but none of them would justify creating a special type of business only for sport from that
organizational perspective.

I have also shown that the exploitation of the market for championship by a league tends to
form itself into a cartel. Cartels are not special; their effect on the markets they exploit has
been studied and is quite clear. It usually results in a suboptimal situation and is not beneficial
for society. Cartels are usually recognized as institutions in which regulation is needed to
protect the consumers and therefore cartels are considered illegal in almost all jurisdictions.
Therefore, cartels are special since they are considered a bad thing for society in general and
must be broken up.

Chapter 1 therefore showed that from an economic point of view, the business of sport is not
special, at least not special enough to justify an exclusion from the application of normal
regulation. In fact it showed us the contrary: since leagues are very close to cartels, they must
be regulated in order to make sure that they do not abuse their position. Then I continued the
analysis since the cartelization could still be a benefit of society in general and there could be
other aspects that are special and could justify the non application of normal laws and
regulations to sport organizations. But this analysis failed to show any benefit from the
cartelisation of sport.

In chapter 2, I examined one of the most important influences on the market for players:
uncertainty of result or, as it is more commonly known, competitive balance.

I showed that competitive balance is a unique aspect of the sport industry because of the
following equation: the closer the contest, the higher the revenues for the participants. For
that, I used the example of boxing, in order to show that the greater the chances of a champion
of being defeated by the challenger, the higher the pay will be for the champion. This can be
explained by the fact that the spectators in sports pay for the excitement and the uncertainty and one can imagine that the higher this uncertainty is, the higher they are willing to pay.

In this chapter, I have shown that competitive balance is one of the most important influences on the demand for sporting contests. Even looking at different aspects of the demand for sport coming from different kinds of fans, I have shown that competitive balance is always an important aspect of the demand and therefore of the revenues of the sport organizations. However, a real definition of the attribute ‘competitive’ remains difficult to find. This is mainly due to the fact that depending on the point of view, there are different ways of measuring competitive balance. Evaluating competitive balance on a game per game basis or over the whole season will result in different conclusion as to the balance in the league. In consequence, I decided to adopt a definition of competitive balance which serves the purpose best and has the most reflection on the market for players. The definition focuses on post-season play and European competition in order to reflect the fact that this will result in higher revenues and more opportunities to influence the value of the players.

At the end of chapter 2, I showed that competitive balance is indeed a special aspect of sport. That aspect may not be sufficient enough to justify a complete exclusion from the laws and regulations applicable by itself, but it may end up being a factor for the specificity of sport in the final analysis.

In chapter 3, I examined the way ‘value’ is determined in the sport business. The goal of this chapter was to try to come to a definition of value in order to have an objective measure, for the purpose of this thesis.

The first step was to start looking at a way to determine the value of the teams and the influence this value has on the value of players. For this purpose, I first examined some of the structures that can be taken by sports teams. This led to the study of the “sportsman owner” theory that shows the particular incentives that are given to the team owners and managers. Since they receive all the benefits of having a successful team but have to pay only part of the costs of what is needed to get that success, the sportsman owner will always spend more than the optimal amount of money for the salary of players and reduce the objective value of the team because of these expenses.
However, the most important conclusion of this exercise was to discover that the value of the teams is linked mainly to one single factor: their share of the monopolistic rent that is exploited by the league. This is where all the wealth of sport organizations comes from and this is what constitutes the value of teams. Again, this makes for a special case generally, however, since we are talking about a monopoly, it is not a desirable situation and there is a need to have the supervision of the law to prevent abuses.

As for the value of players, I settled on what constitutes in my opinion the best way to determine their value: the willingness of club to pay to acquire their services. The simple equation that I used allows me to show the reasoning that a rational team should make when deciding to acquire a player. It also reflects the conflict between the player and the team in which both parties try to get the highest portion of the value of the player for themselves. It was with this equation that I examined the influence of a change of the rules of the game to the value of players in the next chapter.

In chapter 4, I examined the impact of a change in the rules of the game. In order to properly evaluate the impact of these rules, I separated them into two categories: radical rule changes and mild rule changes. I showed that radical rule changes result in the creation of a completely new scale of the determination of value of players. It was difficult to make a before and after comparison since it resembles comparing two different sports.

For the mild rule changes, I have shown that they mainly result in a redistributive effect between the players, assuming that there is no change in the aggregate revenues of the sport. This redistribution operates itself in the following way: players getting an advantage from the change in the rules become more valuable whereas players disadvantaged by the change in the rules become less valuable. It is hard but not impossible for players to predict in which category they will end up, but the main problem is that they do not have much opportunity to adapt since the rules changes are usually done only by a decision of the teams. However, this redistribution only applies to the first generation of players that are affected by the rules. When the second generation of players enters the league, they have already played with these new rules for their entire career so for them it does not constitute a change anymore which leads to a disappearance of the redistributive effects.
The conclusion on the change to the rules of the game is that radical rule changes result in a complete change of the scale under which the value of the player is determined. As for the mild changes in the rules of the game, they result in a temporary redistribution of the value between the players depending on how the individual player is affected by the change. All these situations do not show any special aspect of the sport industry or do not justify preventing the application of the law to sport organisations. This redistribution could be in need of supervision in order to protect some vulnerable parties or to make sure that it does not result in a loss of wealth for society in general.

Concerning the changes in the rules of the commercial game, I decided to compare the restriction to the free movement of players in two sports. This rule was chosen since it is one of the rare rules of the commercial game which is directly aimed at influencing the market for players and should therefore have a clear impact on the value of players. In addition, it is the rule that resulted in most of the court cases implicating sport organisations. It is as a result of the contestation of the restriction on their movement by players that sport organisations claimed to have a special status. The sports that I chose to examine are football and soccer since they offer a very good contrast in the way that the restrictions were decided and implemented.

Football and the NFL showed us an example of a negotiated settlement between the players and the teams on the nature of the restrictions that would affect the freedom of players. After a long string of legal decisions that gradually liberated the players from the severe restrictions that existed in the NFL, both parties seemed to have realized that a negotiated solution might be in their best interest. This was made possible by the special structure of the North American sport market where all players of the top league are represented by a strong union. The consequence was that the players and the teams decided to implement their negotiated solution in the CBA of the league.

The NFL CBA that I studied showed that both parties decided to act as business partners and to decide on the split of revenues of the industry between the parties, the players and the teams. The freedom of movement of players was used as a bargaining chip for each party in order to influence the portion of the revenues that was going to be allocated to the players. The reasoning of the players was that they were ready to have less freedom of movement or have it only later in their career in exchange for a larger percentage of the revenues of the
sport. So they were willing to trade off some individual privilege for the greater benefit of all players. This operates a redistribution of revenue and value between the different categories of players with the players with more freedom being more valuable than the players that are less free to negotiate their employer. Unfair as it may seem, this distribution was agreed upon. The players, through their union, freely negotiated these restrictions and this redistribution with the team. They made the choice of putting this structure in place, knowing most of the impacts on the value of players or at least on the salary of the players. As a business partner of the teams, the players hoped that this system would allow the NFL to generate higher revenues and, since they would have the right to a portion of these revenues, they would also benefit from these revenues. So the current NFL situation is the result of choices made by the parties of the industry. But these choices have been made under the “shadow” of normal law and for the tribunals. The parties know that the redistribution that they agree to could be contested in front of the courts and it forces them to recognise that they are not special and that they cannot ignore normal law.

In soccer, I have shown that the change in the commercial rule is the result of a top down approach. This industry was also faced with a legal decision that gave more freedom to the players but, contrary to the situation of the NFL, the players were not made a partner in the solution that followed the *Bosman* case. The absence of strong collective action by the players resulted in them being excluded from the conception of the rule.

FIFA has always taken the top down approach and it is doubtful if a credible representative players’ organization would have been considered a partner for FIFA even if it had existed. Soccer is a more fragmented market than North American sports and does not have a CBA tradition which makes the FIFA top down approach the only suitable approach so far in the eyes of the industry.

The solution of FIFA also resulted in a redistribution of value between different types of players. In this case we have a twofold redistribution.

First, like in the NFL CBA, we have an age based restriction. This is a restriction to the freedom of movement of a certain category of players. By having the freedom of young players restricted, it redistributes part of their value to the older players that are free to negotiate their contracts with all teams.
Secondly, there is a geographically based restriction. By creating a confederation wide price for training and development of players, the FIFA system redistributes value between players based on their country of origin and the country of the team trying to acquire them.

But, unlike in the case of the NFL, the FIFA system was imposed without negotiation with the players. I also expressed that the whole system leaves the impression that some of the determinations were made in an arbitral manner without much link with reality. In addition, instead of accepting the application of normal law the world of soccer has been lobbying for an exemption to law as recognition of the “special status of sport in society.” The NFL embraced the fact that it is not special but FIFA is still fighting against the fact.

In my evaluation of the impact of a change in the rules of the commercial game, the comparison between both sports allowed to see that the result was in fact the same: a redistribution of value between the different classes of players created by these rules. The difference was how the rules were changed: the NFL chose to negotiate the change, FIFA imposed it. In addition, these rules affect the redistribution of value between the sport organisations and the players. Therefore, the supervision by normal law of these rules is required since the rules are not special but part of a normal business relationship between the players and the sport organisations. In that situation, abuses need to be controlled.

**Final thoughts**

Rules in professional sport have been in constant fluctuation since it has been organized in a structured way. The changes in the rules resulted in many effects - intended and unintended - that in turn affected in different ways the various parties of the industry.

For the value of players, I have shown that pure creation or destruction of value is difficult to achieve only with changes in the rules. Making a player, or players, more valuable in absolute only by changing the rules is a very difficult feat to achieve and to prove.

---

As the situations studied in this thesis show, the most likely impact of a change in the rules will be a redistribution of value between different categories of players. The importance and duration of any redistribution seems to depend on the type of rule changes that lead to it. Therefore sport is not special in that aspect but only reflect the normal situation that is also present in other industries.

However, the most important observation that was made in this work is that the CBAs, the type of commercial rules common in North America, are also a basic exercise in redistribution of value, first between the teams and the players and then between the players themselves. What makes the CBAs particularly interesting is the fact that they are negotiated instead of being imposed; they are the result of a compromise between all the parties concerned.

For future research it would be interesting to concentrate on the different types of CBAs that are present in sport leagues and examine the way in which they are different from the CBAs in other industries. Maybe some special aspect of sport can be found there. Although I failed to find any special aspect of the sport industry in this thesis this does not mean that the search for the specificity of sport will not continue to be an important subject for all interested in the business of professional sports. It may have to move toward very specific aspects and I hope that my humble work helped to open some avenues of research on that subject.
Bibliography

Articles


**Books**


Cases

Aldershot Football Team v. Banks (unreported)


Christelle Deliege vs. Asbl Ligue francophone de judo and others, C-51/96 and C-191/97 (joined)


Clarett v. NFL, United States Court of appeals for the 2nd Circuit, docket 04-0943, May 24 2004


Donà vs. Montero [1976] ECR 1333

Eastham v. Newcastle United [1964] Ch. 413


James McCoy Smith v. Pro Football Inc. and the National Football League [1979] 593 F.2d 1173

Kingaby v. Aston Villa (unreported)

Mackey vs. NFL 407 F.Supp. 1000 (D.Minn 1975), appeal: 543 F.2d 606 (8th Cir. 1976)


Meca-Medina and Majcen vs. Commission, C-519/04

Mid-South Grizzlies v. NFL, 720 F.2d 772 (3d Cir. 1983)


North American Soccer v. NFL 670 F.2d 1249 (2d Cir. 1982)

Powell v. NFL, 690 F.Supp 812 (D. Minn. 1988)


Texaco, Inc. v. Dagher, 2006 WL 461525 (Feb. 28, 2006)

Union Royale Belge des Sociétés de Football Association and Jean-Marc Bosman, C-415/93

Walrave vs. Union Cycliste Internationale [1974] ECR 1405

Conferences


Interviews

Alan Vessey, Football Services Manager of Football Federation Australia. March 2006
Julien Brisebois, Vice-president to Hockey operations of the Montreal Canadien, March 2007

Lawrie Woodman, National Coaching Development Manager & Game Development of the Australian Football League, March 2006

Pedro Trangrouse, Vice-President Legal affairs of the Brazilian Football Federation, November 2005

Stephen Weatherill, Professor, Faculty of Law Oxford University, November 2007

Yves Vonlanthen, Administration & Legal Manager of the International Ice Hockey Federation, March 2007

**Newspaper articles**


FORBES, 2008. The Business of College Football, 13th August


JAMES, S., 2006. Why clubs may risk millions for riches at the end of the rainbow. The Guardian, August 5, available at: http://football.guardian.co.uk/comment/story/0,,1837801,00.html, download date: August 2006


MAHONEY, R., 2007. English soccer fans wary as Yanks take over teams. USA Today, No. 4758, May 16, p.1A


NEW YORK TIMES, 1906. Changes in football rules. August 12

NEW YORK TIMES, 1907. Football’s Death Record for 1907. November 24


RING MAGAZINE, (1938). Louis’ only setback, July 1938


SHUTT, S. 2000, Long live the three. *Naismith Memorial Basketball Hall of Fame*

SWEET, D., 2008. These NFL players are in a class by themselves: With short careers a concern, NFL players look to score at business school. MSNBC, March 5. Available at: http://www.msnbc.msn.com/id/23475390/, date of download March 2008

THE ECONOMIST, 2006. In a league of its own. April 29th, p.63


Public Legislation and reports


Treaty of Lisbon amending the Treaty on European Union and the Treaty establishing the European Community, signed at Lisbon, 13 December 2007, article 2, line 123 to 125

Private Publications


FIFA, Annex to FIFA Circular 959

FIFA, Circular 1085

FIFA, Circular 1142
FIFA Executive Committee, *Regulations for the Status and Transfer of Players*, July 5 2001

FIFA Executive Committee, *Regulations governing the Application of the Regulations for the Status and Transfer of Player*, July 5 2001

FIFA Executive Committee, *Regulations for the Status and Transfer of Players*, December 2004

FIFA Executive Committee, *Commentary on the Regulations for the Status and Transfer of Players*


National Football League, *Collective bargaining agreement*

**Websites**


BBC News: http://news.bbc.co.uk

Eurofootball.biz: http://www.eurofootball.biz


FIFA official website: http://www.fifa.com

FOOTBALL ASSOCIATION, official website. Available at: http://www.thefa.com/
Forbes Magazine: http://www.forbes.com

JOHNSON, P.M., A Glossary of Political Economy Terms by Dr. Paul M. Johnson available on line at: http://www.auburn.edu/~johnspm/gloss/ date of download march 2007

Major League Baseball: http://www.mlb.com/index.jsp

Modern Pentathlon Australia website:  

NFL: http://www.nfl.com

NFL Players Association website: http://www.nflpa.org/

NHL Website: http://www.nhl.com/


The baseball database of Sean Lahman: http://www.baseball1.com