Competitive Balance in Complex Professional Sports Leagues

by

Morten Kringstad

Submitted in accordance with the requirements for the degree of
Doctor of Philosophy

The University of Leeds
Leeds University Business School

June 2008

The candidate confirms that the work submitted is his own and that appropriate credit has been given where reference has been made to the work of others.

Abstract

Competitive balance is seen as crucial to the viability of professional sports leagues, and it has been a central concern of sports economists, industry practitioners and fans. But the concept is complex and ill-defined. The theoretical approach in this thesis is to first analyse competitive balance in a simple league context. It is shown that even in a closed round-robin league with a single prize (i.e. league championship), there are three dimensions of competitive balance: win dispersion, performance persistence and prize concentration. Further extensions to this three-dimensional approach are required when the analysis moves to complex real-world league structures which are typically multi-prize tournaments, with post-season playoffs and often open merit hierarchies with promotion and relegation. A new concept of competitive intensity is introduced.

The three-dimensional approach is applied to the empirical analysis of competitive balance in European (association) football and the North American Major Leagues. It is found that cross-league comparisons of competitive balance are dependent on the dimension analysed. Win dispersion is better in the Big Five European domestic football leagues compared to the Major Leagues, but the reverse is found for championship concentration. The Major Leagues are also found to be more competitive in respect of the concentration of post-season playoff qualification. Differences in competitive balance are found between the European domestic football leagues. The causes of these cross-league differences are investigated with regression analysis. Again the results are highly dependent on the specific dimension of competitive balance. Win dispersion is significantly associated with national geographic and economic characteristics as well as league structure. Time-series analysis is undertaken to study changes over time in competitive balance in the top divisions of the Norwegian and English football leagues. The results are most significant for the English league, and run counter to the predictions of the invariance principle.

Acknowledgements

I am grateful to a number of persons who have been helping and supporting me during the process of writing this thesis. The thesis is part of the research program "Analyses of economical consequences of sports at top level - where football, media and events are emphasised" at Trondheim Business School, a faculty of Sør-Trøndelag University College in Norway, which was a participant of the programme "Sport, Society and voluntary sector", at the Research Council of Norway. Hence, I am grateful to The Research Consul of Norway for financially supporting this thesis, and to programme coordinator Tore Abrahamsen, who has always been helpful. Also thanks to Solbjørg Rauset and Jane Hildebrand at the Research Consul of Norway for their administrative support. I am grateful to my colleague Harry Arne Solberg, Professor at Trondheim Business School, who inspired me to start on this thesis, and has always been supportive. He has also been the coordinator of the research programme mentioned above at Trondheim Business School. Also the support, both financially and through flexible work commitments, from Dean Ove Gustafsson at Trondheim Business School, has been very important. The many meetings with my colleague Hilde Fjellvær throughout the writing of this thesis has meant a lot to me.

Many thanks are due to psychology student and my sister-in-law Anette Hassel, who has helped me typing raw data of league tables, as well as typing information into a reference software program. Research assistants at Trondheim Business School Tor-Eirik Olsen and Baard Dahl Svendsen have given me a helping hand by computing the Spearman's rank correlation coefficient over several leagues. Thanks are due to them and to Dag Vestlund at the Norwegian football association and Boye Skistad at Norsk Toppfotball (NTF) for informing me about the history of Norwegian football, as well as to PhD student Hallvard W. Johnsen for reading through parts of Chapter 6. I am also grateful to my wife, Tone, for correcting my English several times and to Hilde Stuve Kjellsmoen for language assistance at the end of the writing up period. Also thanks to Bente T. Hansen for advices on language.

At the University of Leeds, a lot of thanks go to Sarah Gillings and Natasha Mullea at the Research Office of Leeds University Business School for minimising administrative problems for a split site student. Also thanks to Dr. Tom F. Burgess for good advice. Not least, a very special debt is owed to my supervisor Professor Dr Bill Gerrard at

Leeds University Business School. It has been a great experience to be supervised by such a capacity in the field of the Sports business, economics and management, who is also very friendly on the personal plan. During my stays in Leeds, a lot of hospitality has also been given me from his wife, Jackie, and their daughters, Anna, Isla and Laura.

Finally and not least, I want to give my wife, Tone, and my children, Thea, Julia and Aurora a lot of love for their support and patience. Also many thanks are owed to the rest of my family for their support.

Own Publications and Presentations Relevant to this Thesis

Chapters in books and not-refereed publications:

Kringstad, M. (2003). Konkurransebalansen i Norsk og Europeisk Fotball. 2003. TØH-serien nr 2003:1.

Kringstad, M. (2003). Konkurransebalansen i Norsk og Europeisk Fotball, In Fallan. L. og Gustafsson, O. (Eds.), *Frihet og mangfold. Festskrift til Odd G. Arntzen.* Trondheim, Norway: Trondheim Økonomiske Høgskole.

Kringstad, M. & Gerrard, B. (2004). The concepts competitive balance and uncertainty of outcome. In G.T. Papanikos (Ed.), *The Economics and Management of Mega Athletic Events: Olympic Games, Professional Sports and Other Essays.* Athens, Greece: ATINER.

Kringstad, M. & Gerrard, B. (2007). Beyond Competitive Balance. In M. Parent and T. Slack (Eds.), *International Perspectives on the Management of Sport.* Burlington, MA, San Diego, CA and London, UK: Elsevier.

Conference presentations:

Kringstad, M. (2002). *Competitive Balance in Norwegian and European football.* 2002 European Association of Sport Management (EASM) Conference, September 4-7, 2002, Jyväskylä, Finland.

Kringstad, M. & Johnsen, H. (2003): *Demand for Norwegian and Swedish football.* 2003 EASM Conference, September 10-13, 2003, Stockholm, Sweden.

Kringstad, M. & Gerrard, B. (2004). *Competitive Balance in European Football (Soccer).* 2004 EASM Conference, September 22-25, 2004, Ghent, Belgium.

Kringstad, M. & Gerrard, B. (2005). *Differences in Competitive Balance across Leagues in European Soccer.* 2005 North American Society for Sport Management (NASSM) Conference, June 1-4, 2005, Regina, Saskatchewan, Canada.

Kringstad, M. & Gerrard, B. (2005). *Theory and Evidence on Competitive Intensity in European Soccer.* 2005 International Association of Sports Economists (IASE) Conference, July 18-19 2005, Ottawa, Canada.

Kringstad, M. & Gerrard, B. (2005). *Competitive Intensity in European Football.* 2005 EASM Conference, September 7-10, 2005, Newcastle, England.

Kringstad, Morten (2006). *Consequences of UEFA Champions League in Small Leagues.* Short presentation at the UEFA/IASE Event: "The economic impact of the UEFA Champions League on National Football Economics", May 2, 2006, Nyon, Switzerland.

Kringstad, M. & Gerrard, B. (2006). *Competitive Balance across European Soccer Leagues.* 2006 IASE Conference, May 4-6, 2006, Bochum, Germany.

Kringstad, M. & Gerrard, B. (2006). *Prizes and Tournament Structure in Professional Teams Sports.* 2006 NASSM Conference, May 31-June 3, 2006, Kansas City, Missouri, USA.

Kringstad, M. & Gerrard, B. (2006). *Competitive Balance in North American Major Leagues and European soccer Leagues.* Part of the symposium: Cash, Caps and Coaching: An International Perspective on the Drivers of Competitiveness and Competitive Balance in Professional Sports Leagues. 2006 NASSM Conference, May 31-June 3, 2006, Kansas City, Missouri, USA.

Kringstad, M. & Gerrard, B. (2007). *Competitive Balance in a Modern League Structure.* 2007 NASSM Conference, May 30-June 2, 2007, Ft. Lauderdale, Florida, USA.

Kringstad, M. & Gerrard, B. (2007). *Competitive Balance and Free Agency in the English Top Division.* 2007 NASSM Conference, May 30-June 2, 2007, Ft. Lauderdale, Florida, USA.

Kringstad, M. (2007). *Competitive Balance in European Football.* 12th Annual Congress of the European College of Sport Science, July 11-14, 2007, Jyväskylä, Finland.

Newspaper article:

Kringstad, M. (2007). Competitive Intensity and the Interest for the English FA Premier League. *Adresseavisen* (Norwegian newspaper, located in Trondheim), September 2006. [It is written in Norwegian].

Table of Contents

List of Tables

List of Figures

1. Introduction

Despite competitive balance being central in the academic literature, as well as among industry practitioners and sports fans, it is still, more than half a century after the seminal paper in the field of the economics of professional team sports by Rottenberg (1956), an ill-defined concept, both theoretically and empirically. This is evidenced by the large number of measures trying to quantify the level of competitive balance. As Gerrard (2004b, p. 45) comments, "*competitive balance* has become a catchall for the different aspects of outcome uncertainty and contest significance". Confusion about the conceptual meaning of competitive balance undermines the understanding of the relationship between relative competitiveness in sporting contests/tournaments and economic value. Contributions that help create a common base for the concept are therefore important, and may help resolve some of the key political issues within professional sports league over the regulation of labour and product markets to promote greater competitive balance

This thesis concentrates on issues related to the concepts of uncertainty of outcome and competitive balance in professional team sports. The principal focus is European professional (men's) football, with special attention to the top divisions in Norway and England. Norway is a small country in population terms, with only close to 4.7 million citizens.[1] As a consequence, the Norwegian Football League is small compared to many other European football leagues. On the other hand, the English FA Premier League is probably the most popular football league internationally. The changing structures in European football during the 1990s have raised many questions regarding uncertainty of outcome and competitive balance in leagues in an open international market.

1.1. Analytical Framework, Research Questions and Limitations

To be able to afford player wages, the fundamental cost in professional team sports, teams must generate sufficient revenues. In professional team sports, there are a number of income sources, which have increased in scope and scale during the last thirty years

[1] According to Statistics Norway, the number was 4,681,134 at the 1st of January 2007 (www.ssb.no/emner/02 01 10 folkemengde/tab-2007-03-08-01.html).

in European football.[2] However, gate attendances can still provide the best indicator of a team's revenue potential, not only matchday revenues, but also other revenue streams (i.e. media, sponsorship, merchandising and other commercial income). Understanding the underlying factors for gate attendance demand is hence important, as described in Peel and Thomas (1988, p. 242):

> The nature of the demand function in any professional team sport is important for the analysis of club and league behaviour and decision making. An understanding of the determinants of demand should influence decisions on price, investment in both physical and human resources, marketing and product characteristics - which has implications for league structure and organisation, and competition rules.

The demand for professional team sports has been analysed, both in European football and in other leagues, and is also touched on by Rottenberg (1956).[3] There are a number of drivers for demand in professional team sports. Some of these might be categorized as sports marketing drivers. However, this thesis will concentrate on sports economics drivers. Peel and Thomas (1988) divide the determinants of gate attendance demand into four groups; economic variables, geographic and demographic variables, variables related to the attractiveness of the fixture, and determinants "which can influence attendance at a match on the day" (p. 245). Absolute and relative team quality are relevant factors in the group of variables related to the attractiveness of the fixture. This thesis will focus on the first, relative quality, which is connected to the concepts of *uncertainty of outcome* and *competitive balance.* Hence the conceptualised demand function takes the following form:

Demand = f(uncertainty of outcome, other factors)

Uncertainty of outcome is hypothesised to be one of the factors driving demand for professional team sports. Further, it is expected that closer competition between the teams in a league, which means better competitive balance, will increase the uncertainty

[2] Andreff and Staudohar (2002) explain this by comparing two models, where the SSSL model (Spectators-Subsidies-Sponsors-Local, p. 25) reflects the traditional business of European team sports, while the MCMMG model (Media-Corporations-Merchandising-Markets-Globalization, p. 32-33) is based on modern business, particularly in big leagues.
[3] See, for example, Cairns et al. (1986), Downward and Dawson (2000), Dobson and Goddard (2001), and Borland and Macdonald (2003) for overviews.

of outcome (see, for example, Rottenberg, 1956). In other words, demand is expected to be affected by competitive balance through uncertainty of outcome, as is argued for example in Scully (1989), who claims that the main determinant for uncertainty of outcome is the competitive balance among the teams:

Uncertainty of outcome = f(competitive balance, other factors)

Because of this hypothesised relationship between competitive balance and the demand for professional team sports through uncertainty of outcome, competitive balance has been central and regarded as crucial to the economic well-being of professional team sports. This importance is not only related to the maximisation of demand (and revenues) by individual teams, but also at the collective level there may be a need to maintain a minimum level of competitive balance in order to ensure the viability for leagues.

Even although competitive balance is a frequently used term in the literature of business and economics of professional team sports (see, for example, Fort and Maxcy, 2003; Fort and Quirk, 1995), a common unique definition of competitive balance in this literature is missing (Kahane, 2003). This is, for example, emphasised in Szymanski (2003a, p. 471): "One problem that arises is the precise definition of competitive balance", and this is reflected in the number of terms and explanations related to the concept of competitive balance that can be found in the literature. Competitive balance has been described[4] variously as the differences in playing strengths between the teams in a league (see, for example, Scully, 1989), the distribution of playing strengths (see, for example, Fort and Quirk, 1995; Dobson and Goddard, 1996; Késenne, 1996, 2002a), the comparable playing strength of teams (see, for example, Bennett and Fizel 1995; Kuypers 1997; Sutter and Winkler 2003), the equalisation of playing strengths (see, for example, El-Hodiri and Quirk 1971; Scully 1989; Forrest and Simmons 2002), the balance of playing strengths among teams (see, for example, Rottenberg, 1956), and the relative strengths among competitors (Sanderson and Siegfried, 2003). Describing competitive balance in terms of the playing strengths of teams tends to imply a concern directly with match and/or tournament outcomes, either actual or expected. Competitive balance is also referred to as quality differences between the teams with respect to the

[4] These are based on my interpretations, and much of what I have called descriptions of competitive balance in the literature has not been done explicitly.

relative quality of team play (see, for example, Demmert, 1973; Scully, 1989; Vrooman, 1996; Cyrenne, 2001) and the more or less equal quality of play among teams (see, for example, Rottenberg, 1956). In the case of quality of play the emphasis, at least implicitly, is on the components of the team performance rather than necessarily the outcome. A popular description of competitive balance is as the distribution of playing talent (see, for example, Rottenberg, 1956; Sloane, 1969; Vrooman, 1995, 1996, 2000; Késenne, 1996, 1999; Hausman and Leonard, 1997; Szymanski and Kuypers, 1999; Dobson and Goddard, 2001; Depken, 2002; Leeds and von Allmen, 2002; Bourgheas and Downward, 2003; Fort, 2003). In this case, the focus is on the input to the sporting production function rather than the intermediate and/or final outputs.

In general, competitive balance has been related to and reflected from many angles when it comes to competitions. Zimbalist (2003a) describes the concept as both complex and multi-dimensional, as well as being reflected by both historical and cultural aspects of a sport. This has made it almost impossible to come up with a simple, generally acceptable definition in the literature. However, this thesis is motivated by the importance of having a common understanding of competitive balance both theoretically and empirically.

The suggested definition of competitive balance in this thesis will be based on simplicity. A generally accepted conceptual framework for competitive balance should also include definitions of related concepts to better distinguish between them. This is desirable in order to facilitate comparisons between different empirical studies which may arrive at different conclusions in part because of their use of different notions of competitive balance. It is also important that the conceptual framework is robust to different levels of complexity in league structures.

Treating competitive balance mentioned in terms of sporting quality yields a definition of competitive balance as *the distribution of sporting quality between the teams in a league/tournament*. Such a definition may be appropriate in those leagues using a traditional (balanced) round-robin match schedule. But is it still a robust definition in more complex league structure with, for example, unbalanced schedules that depart from a pure round-robin format? This thesis will instead use the description above as a definition of what can be seen as a related expression to competitive balance, namely, the *sporting quality balance* (QB).

If the teams in a league play against each other the same number of times, and an equal number of home and away matches, the league can be said to have a balanced match schedule. Currently match schedules in a number of leagues are unbalanced. This might affect the distribution of sporting outcome in those leagues such that the expected distributions of sporting quality and sporting outcome might not be identical. Since it is desirable that a competitive balance definition is also robust to different systems of match schedules, this thesis suggests using the sporting outcome as the basis for the construction of a general definition of competitive balance. *Competitive balance is the distribution of sporting outcomes in a league/tournament.* However, since the sporting quality balance is the main (direct) component in competitive balance, but also because the match schedule might "disturb" this relationship, the following can be written:

Competitive balance = f(QB, match schedule, others)

This suggested definition of competitive balance might be better able to capture many other descriptions of competitive balance found in the literature, such as competitive equality (see, for example, Jones, 1969; Demmert, 1973; Daly and Moore, 1981; Cairns et al., 1986; Jones and Ferguson, 1988), balanced competition (see, for example, Cairns et al., 1986; Késenne, 2002b; Vrooman, 2000), closeness of competition (see, for example, Davenport, 1969; Sloane, 1971; Daly and Moore, 1981; Vrooman, 2000), "tighter" competition (Rottenberg, 1956), equality among clubs (see, for example, Sloane, 1969), sporting equality[5] (Dabscheck, 1975a, 1975b; Sloane, 1976b), equality of competition (Dabscheck, 1975a; Sloane, 1976a), parity[6] (see, for example, Depken 1999; Sutter and Winkler, 2003; Cain and Haddock, 2006; Fort, 2007), and the evenness of sporting competition (Macdonald and Borland, 2004). Noll (2003, p. 549) claims that "competitive balance refers to the extent to which teams are sufficiently closely matched that game outcomes are uncertain."

Further, the proposed definition of competitive balance takes into account the essence of the meaning by competitive balance, often expressed in various measures, based on the distribution of sporting outcomes between the teams in a league/tournament, as is emphasised in Humphreys (2003, p. 286):

---

[5] This expression might be related to the (sporting) quality distribution.

[6] This is a general expression. If it is based on sporting outcome, then it is appropriate in the context of this thesis.

Sports economists and others have developed a number of different measures of competitive balance. All measure the distribution of wins and losses across teams in sports leagues; each represents a different method of describing this distribution.

The definition of competitive balance in this thesis takes a more restricted approach than the general description of competitive balance, as "a catchall term that refers to a number of different aspects of competition on the playing field" (Quirk and Fort, 1992, p. 244).[7] The complex and ill-defined nature of the concept of competitive balance is related to complications in at least three of the core elements of the economics of professional team sports.

First, identifying appropriate measures to quantify the level of competitive balance of a league requires a clear understanding of what needs to be measured. The understanding of competitive balance affects how it is to be measured and vice versa. The complexity of the concept is reflected in the wide range of measures applied in the literature (see, for example, Fort and Maxcy, 2003).

The second element is based on the relationship, discussed earlier, that competitive balance is hypothesised to be an indirect demand driver. Before evidence from analyses of these relationships can be seen as reliable, it is required that the variable(s) reflecting competitive balance is (are) valid. The relationship between competitive balance and uncertainty of outcome is entirely dependent on restricting of the concept of competitive balance. The wider the use of competitive balance as a catchall term, the more difficulty it is to find a clear relationship between these two concepts. They become synonymous. Understanding the hypothesised relationship between competitive balance and demand through uncertainty of outcome therefore not only requires an understanding of the limitations of competitive balance as a concept, but also that of uncertainty of outcome. Like competitive balance, uncertainty of outcome is complex and relates to different dimensions of the consumption decision. Mapping similarities and differences are important in understanding the hypothesised relationships, and also for being able to structure appropriate empirical analyses to measure the significance of this relationship.

[7] This is followed by Bennett and Fizel (1995).

This thesis aims to develop such a mapping, as well as clearly identifying the concept of competitive balance separately from other competitiveness-related concepts.

Third, complexity is not only directly related to the concept of competitive balance, but also to that of the intended target, the tournament. Multi-prize tournaments are hypothesised to have diverse impacts on the relationship between consumer demand and sporting competition over time. Leagues are complex. Hence, complexity in relation to the competitive balance is not only related to the concept itself, but also to the complexity of league structures. The combined complexity complicates the landscape for measures of competitive balance. This thesis aims both to capture the relationship between complexity in competitive balance as well as in tournaments, and, in addition, take advantage of both the complexity of real-world league structures and the diversity of measures of competitive balance, to show that a portfolio of indicators must be used to assess competitive balance levels across complex, real-world leagues.

The concern of this thesis is not just related to measuring the level of competitive balance for a specific league in a single sport. A better understanding of competitive balance is acquired by comparing competitive balance across leagues and across sports. A natural follow-up is to analyse the determinants of competitive balance both by cross-sectional studies across leagues as well as time-series analysis of individual leagues. This thesis aims to deal with all of these empirical issues.

Part of the complexity of the concept of competitive balance is related to its multi-dimensionality. This thesis will focus on the three general dimensions: *within season competitive balance, across season competitive balance* and *prize concentration*. The suggested definition of competitive balance in this thesis is robust to these three dimensions, and makes it possible to highlight them, because they are all important parts of the competitive balance concept. How to treat competitive balance in a simple way requires recognising the three dimensions, but not to over-complicate the relationship between them. In other words, this thesis does not attempt to formulate the ultimate measure of competitive balance, but instead recognises the three dimensions, and treats and measures them as important parts of the total picture of competitive balance. It will be shown that the three dimensions are also valid in a simple league context. A number of the measures used in the literature can be argued to be related to a simple league context, and some of them have to be adjusted in order to be appropriate for complex real-world league tournaments. This thesis aims to categorise different measures into different dimensions of competitive balance, as well capture the real-world complications.

This thesis will have a different approach to many of the problems mentioned above compared to the previous literature, by focusing on the concept from a simplified point of view first before turning attention to complexities. This argument turns into the main research question of this thesis: *Can the understanding of the complexity of competitive balance and related issues be improved by focusing on their simplicities?* The main research question can be further split into a number of research questions about the phenomenon of competitive balance, such as what does the expression really mean? What are the theoretical implications of competitive balance? How should competitive balance be measured? And what are the determinants for competitive balance?

Some of the issues involving competitive balance can be illustrated by using the Norwegian football league as a small case study. One team has dominated this league in recent years, winning every single league championship between 1992 and 2004. This team, Rosenborg BK, has also been in a "league of its own" regarding financial strength, mainly because of substantial additional revenues earned from qualification to the UEFA Champions League, but also from developing players that have been sold abroad profitability (at least until 2000).[8] This success is particularly noteworthy since Rosenborg BK is located in only the third largest market in the Norwegian league.[9]

The domination of a sports league by one team, as has been in the case in the Norwegian top division in recent years, has been frequently discussed in the literature on the economics and business of professional team sports. The Louis-Schmelling paradox (Neale, 1964) emphasises that being too strong compared to (the) competitors can result in the league as a whole diverging significantly from profit maximization, and at the extreme Neale states that in sports: "Pure monopoly is disaster" (p. 2). In professional team sports the New York Yankees is used as a similar case, called the Yankees paradox (Vrooman, 1996).[10] It is related to what happened at the end of the

[8] This can for example be found in annual reports and several newspaper articles.
[9] My anticipation related to the size of the city (see, for example, www.ssb.no/emner/02/01/10/folkber/tab-2006-12-20-01.html). However, probably because of the team's sporting success, it has usually been the most watched team in this period, such as in the 2005 season (see fotball.vg.no/bors/tilskuer_klubb.php?sesong_id=143).
[10] See also Neale (1964) p. 2.

team's "dynasty period" in the 1950s, when the Yankees' gate attendances surprisingly rose in a season that the team did not perform as well as before. These examples can be directly related to the concepts of *uncertainty of outcome* and *competitive balance*.

The dominating position of Rosenborg BK in modern Norwegian football began at the appointment of a charismatic coach in 1988. Apart from his second and fourth seasons (when the team finished as runners-up), Rosenborg BK won every championship until he retired after the 2002 season.[11] If Rosenborg BK in the Norwegian league followed the Yankee paradox, one would expect the attendance to have risen in the 2005 season when the team lost its dominating position and finished only seventh in the league. Below, table 1.1 shows average attendance and the point difference between Rosenborg BK and the best placed rival team since 1988.

Table 1.1: Rosenborg's average attendance in the period 1988 – 2005 and the point score difference to the winner/second team

Season	Att.	ΔPoints	Season	Att.	ΔPoints	Season	Att.	ΔPoints
1988	12 070	7	1994	11 061	8	2000	11 944	7
1989	12 065	- 8	1995	10 280	15	2001	12 070	1
1990	11 115	2	1996	11 062	13	2002	14 626	6
1991	11 451	- 5	1997	11 338	11	2003	15 825	14
1992	13 569	6	1998	13 163	9	2004	17 395	0
1993	11 750	2	1999	13 359	6	2005	17 541	- 12

Sources: www.rbkweb.no/statistikk/snitt.shtml and www.rsssf.com.

Attendance at Rosenborg's home matches over the period shown in table 1.1 and the point difference has a correlation coefficient of -0.343, and is only weakly significant (one-tailed test). This might indicate some positive relationship between attendance at the team's home matches and closeness of the championship race. However, the increased attendance at the end of the period is also related to stadium rebuilding. Therefore, including point difference and rebuilding phase one (from 1995) and phase two (from 2002) as determinants in a regression analysis where attendance is the dependent variable, is more appropriate, and the results are presented in table 1.2.

[11] He had a "one season break" in 1998 (www.rbkweb.no statistikk trenere.shtml).

Table 1.2: Regression analysis on Rosenborg's attendance over the period 1988 – 2005

	Constant	Pointdiff	Stadium1	Stadium2
Coefficient	11983.44	-66.92	497.31	3999.84
Standard error	387.65	37.73	603.99	684.60
P-value	0.000	0.098	0.424	0.000

F-value	Adjusted R^2	DW
21.297***	0.782	1.720

*** Significant on one percent level.

It can be seen that the last rebuilding (full seat) has had a strong significant positive effect on attendance, while, from a competitive balance point of view, it is interesting that "pointdiff" is (weakly) significant. This means that increased point difference, compared to the best placed rival team has a negative response on the number of spectators, ceteris paribus.[12] However, the reduction in the number of spectators, due to higher point difference, is not very high at about 67 spectators on average per point.

The last observation (2005) in table 1.1 showing increased attendance, even in a poor performance season, is interesting in the light of the Yankee paradox, and might reflect that spectators value close competition and/or alternation at the top of the league. Another factor in the 2005 season that might have affected the attendance is that, even if Rosenborg BK did not compete for the prizes at the top of the league, it still had to focus on another sporting prize. This prize, staying in the division and not being relegated, is significant in European football. Rosenborg BK only secured its place in the Norwegian top division with one home match to go. In other words, even if Rosenborg BK performed relatively bad in the 2005 season, many of the matches were significant and went into what Jennett (1984) calls "relegation significance". The relegation significance is based on the hypothesis that relegation battles are more interesting for spectators than matches between middle-of-the-table teams.

The 2004-season also had many significant matches, but this time it was related to the highest valued prize in the tournament, the championship. So even if Rosenborg BK had been dominating for a longer period, the championship in 2004 was highly competitive. The championship was only decided at the very end of the last round of matches in that season (on goal difference). The tight fight for the championship should indicate high

[12] This result seems to be robust, as can be shown by including a season variable and leave stadium1. Including both of these gives high value of VIF (variance inflation factor) and indicates multicolinearity (see, for example, Greene, 2008).

level of "championship significant" matches, which is, according to Jennett (1984), a significant interest driver.

1.2. Outline of the Thesis

To be able to understand and handle the complexity of competitive balance better, Chapter 2 will address a framework that is based on a simple league tournament context. What will be shown in this thesis is that even in a simple league context, competitive balance will comprise three dimensions. This requires a wide set of statistical measures, in order to be able to give a full picture of competitive balance in a league. It can also be argued that the existing literature concerning competitive balance is, to a large degree, based on a simple league structure.

The real-world complexity of sports leagues is taken into account in Chapter 3. Multi-prize tournaments typically deviate from a simple league context, and will demand an extension or modification of measures for competitive balance. Based on multi-prize structures of leagues, it can be seen as appropriate to conceptualise the relationship between team competitiveness and fan interest in terms of the *competitive intensity* of leagues. The basic idea behind competitive intensity goes back to Jennett's (1984) championship- and relegation significance concepts. These concepts have traditionally been related to the uncertainty of outcome.[13] Therefore, competitive intensity will be linked to the basic concepts in this thesis; competitive balance and uncertainty of outcome. Knowledge about issues related to complex league structure and competitiveness within a league in relation to competitive balance can increase the understanding of the consequences of regulatory policies on leagues, and is hence relevant for governing bodies.

Chapter 4 follows up the conceptual framework in the two previous chapters, and will apply empirical measures to categorise the level of competitive balance among European football leagues and North American major leagues (NAML[14]) in respect of the three basic dimensions of competitive balance. While NAML typically have weaker competitive balance in one dimension compared to European football leagues, the opposite usually holds for the second, and the third is relatively equal. Also between

[13] See, for example, Cairns et al. (1986) and Downward and Dawson (2000).
[14] This thesis will use the letters NAML as notation for the North American major leagues, and is hence not following the notation NAL in Fort and Lee (2007).

European football leagues, important differences can be found between different groups of leagues. In addition, measures related to complex league systems will be applied in selected leagues, including Australian leagues. One of these relates to qualification for post-seasonal playoffs, where the empirical findings are somewhat at odds with the general impression of the level of competitive balance between the major leagues in North America.

Chapter 5 concentrates on why differences in competitive balance occur between leagues. A number of determinants will be analysed, with the focus on the European football leagues. Because of structural differences between these leagues, compared to the differences between the NAML, it is argued that the determinants applied on the other side of the Atlantic cannot necessarily be similarly applied in European football. These are specifically related to two groups: (i) factors affecting drawing power; and (ii) prize and tournament structure. The latter is chosen as a consequence of divergence among the European football leagues, as will also be highlighted. Cross-sectional regression analyses show that significance among the determinants is dependent on the dimension of competitive balance to be measured.

Chapter 6 focuses on relationships between competitive balance and changes in regulations and prize and tournament structure, within single leagues over time. Two cases will be presented: the English top division and the Norwegian top division in football. Analyses of these two leagues will, among others, be undertaken using time-series analysis. In general, the response on competitive balance from regulatory changes is much more significant in the English top division than in the Norwegian league. Some of these results are interesting, because they depart from the theoretical invariance proposition. An example is the negative impact on competitive balance in England, after labour market restrictions were relaxed at the beginning of the 1960s.

2. Competitive Balance in a Simple League Context

2.1. Introduction

The concept of competitive balance will be the focus of this chapter. The importance of competitive balance is related to the uncertainty of outcome hypothesis proposed by Rottenberg (1956) that higher uncertainty of outcome increases the interest for matches and the league, ceteris paribus. If improved competitive balance increases the uncertainty of outcome and this is valued by fans, the hypothesised relationship implies that a minimum level of competitive balance is required for a league to be financially viable.

The interest in the concept of competitive balance is not only related to the more fundamental economic and financial consequences. Even today, more than half of a century after Rottenberg's seminal paper, the meaning of the concept is still confused and uncertain. The question what is really meant by competitive balance has been a recurring theme in the literature. The concept is both complex and multidimensional. This is not surprising, since it is related to competition and competitiveness, both concepts that are themselves difficult to define precisely and comprehensively beyond vague and general descriptions. However, it is not satisfactory for this research area to be without common acceptable definitions of such a central concept.

This thesis will take a different angle than the existing literature by considering the concept of competitive balance in different league structures. Many competitive balance measures might be most useful in a theoretical simple league tournament context, while leagues in reality have a higher level of complexity. As a consequence, competitive balance measures may need to be adjusted to take account of more realistic complex tournament structure.

This chapter aims to define what is meant by a simple league context and to give an overview of measures of competitive balance used in the literature. It will concentrate on the conceptual framework of the concept of competitive balance, as well as the related concept of uncertainty of outcome. This framework will include suggested definitions, different time dimensions, relationships between the concepts, and their determinants.

2.2. Central Concept

Competitive sports are organised as contests defined as individual matches between athletes or teams, and as tournaments consisting of a set of contests between a group of athletes or teams. A team sports tournament is, in general, either played as a knock-out/elimination (i.e. cup) competition and/or a round-robin (i.e. league) tournament. This thesis will focus on the latter, where teams play a number of matches, and at the end of the season, the results are aggregated into an end-of-season table. However, many of the issues considered are also relevant to other tournaments. There are, for example, a number of similarities between a round-robin league tournament and a competition such as the Tour de France. Each match and each contest has its winner, in addition to an overall tournament winner. Both modern round-robin team sports leagues and Tour de France are "multi-prize tournaments", where the teams/athletes compete for more than just one overall prize. Examples of other prizes include qualifying for post-season play-offs and avoiding relegation. In the Tour de France there are four major individual prizes - the yellow jersey, the green jersey, the red polka dot jersey, and the white jersey (www.letour.fr).[15] This thesis also uses the term *professional* to indicate the role of financial resources and incentives on the sporting performance for teams. However, many of the results from the analysis will also be valid for amateur sports.

The primary product in professional team sports, the match, is unique since it *cannot* be produced by a single team.[16] In other words, combinations of inputs (teams) are required to produce a match (see, for example, Neale, 1964; Demmert, 1973; Vrooman, 1996). Neale (1964, p. 2) calls this phenomenon the "inverted joint product", or the "product joint". A match requires two teams, and a league tournament needs a several number of teams. In general, this means that one particular team will not survive in a league when the other teams resign (and are not replaced) (see, for example, Jones, 1969). This is a peculiarity of professional team sports,[17] with no parallel in other industries, and first highlighted by Rottenberg (1956, p. 254):

[15] More about Tour de France in a sport economics context can be found in Torgler (2007).
[16] Rottenberg (1956) describes the product as the match weighted by the revenues it is generating.
[17] See also Zimbalist (2003b, p. 503). More about peculiarities of professional team sports can be found in Neale (1964) and Gerrard (2006b). See also Noll (2006) for discussions of Neal's article related to peculiarities.

> If a seller of shoes is able to capture the market and to cause other sellers of shoes to suffer losses and withdraw, the surviving competition is a clear gainer. But in baseball no team can be successful unless its competitors also survive and prosper sufficiently so that the differences in the quality of play among teams are not "too great".

In other words, there is a need for co-operation between the competitors in a league (and a match), or what Gerrard (2000c, p. 202) calls "co-operative competition". The teams therefore have both common collective interests and their own private interests (Knowles et al., 1992; Szymanski and Kuypers, 1999). The individual private interest of teams is to be competitively successful against rival teams. However, between the competing teams, there is co-operation on match schedules and tournament/match rules. These practical issues are decided indirectly by the teams through the league's governing body. The league's governing body represents the teams and is responsible for protecting both co-operation and competition by making and monitoring rules within a framework given by the sport's own international governing bodies (e.g. UEFA and FIFA)[18] and domestic and international laws (e.g. the EU). These rules are not only "sporting rules", but also rules affecting the industry's economy (see, for example, Sanderson, 2002).

The differences between individual and collective interest create an economic dilemma: individual team success versus the negative collective consequences from dominated, predictable leagues. Gerrard (1999, p. 274) claims that this can undermine collective interests: "If the leagues become predictable, fans lose interest and all teams lose financially". Rottenberg (1956) shows this relationship as the main reason for the collapse of the first professional baseball league in the middle of the 1870s. Both Rottenberg and Szymanski and Zimbalist (2005) emphasise that the next professional baseball league to be formed (i.e. the National League) clearly recognised the importance of competitive balance. As a consequence, the so-called reserve clause was introduced in 1879. Moreover, Horowitz (1997, p. 373) claims that: "maintaining league balance in the major leagues was an expressed goal of the 1903 National Agreement that provided the blueprint for the modern major leagues (Davis, 1974, p. 363)".

[18] The Union of European Football Associations and the Féderation Internationale de Football Association.

A league tournament has a number of characteristics (Demmert, 1973; Gerrard, 2004b), and uncertainty of outcome is one of the key elements in the product that professional team sport is offering its customers.[19] Rottenberg (1956) expresses the term "uncertainty of outcome" explicitly, and it is used in arguments such as that "uncertainty of outcome is necessary if the consumer is to be willing to pay admission to the game" (p. 246). The importance of the concept of uncertainty of outcome is therefore built on the hypothesis that a particular match will be more interesting the smaller are the differences in sporting quality between the teams. Since a league consists of many teams, high uncertainty of outcome of the matches played requires small differences in sporting quality between the competing teams. In other words, for the unpredictability matches to be as high as possible, competitive balance must be high. Zimbalist (2003b, p. 503) summarizes these relationships in the following way:

> The success of a league is, to some extent, affected by the degree of uncertainty of outcome of its contests and its seasonal competitions, or, stated differently, by the degree of balance among its teams.

This follows the hypothesis by Rottenberg (1956, p. 7) that "the 'tighter' the competition the larger the attendance". Downward and Dawson (2000, p. 21) explain the uncertainty of outcome hypothesis in the following way:

> This 'uncertainty of outcome' hypothesis argues that close competition between, for example, team A and team B confers benefits on the league not only via increased attendance at teams' own games but also at games involving teams C and D, hence the externality.

The importance of competitive balance is therefore related to the uncertainty of outcome hypothesis on the basis that better competitive balance will increase uncertainty of outcome, and hence be positively related to interest and attendance. Both the match attendee and the television viewer are viewed as valuing uncertainty of outcome (Kuypers, 1997). Actually, Szymanski (2003c) claims that competitive balance is more important for TV viewers, since many of them are more interested in playing quality

[19] Customers are used here to put the product into an economic context. However, they would usually rather prefer to be called supporters and fans.

and uncertainty of outcome than following one particular team. Following Szymanski, the uncertainty of outcome hypothesis might be stronger for TV viewers.[20]

Since it is the interest in the matches, through people at the stadium or people watching television, which usually is the main revenue driver for teams, the uncertainty of outcome hypothesis links competitive balance directly with the profit-earning capabilities of the individual teams (Knowles et al., 1992). Therefore, in general, competitive balance and uncertainty of outcome can be seen as important drivers for the sporting and financial performance of the teams. Since a league is a function of the teams, these factors will also be fundamental at league level. Hence, Zimbalist (2002, p. 119) puts it, "fan perceptions and behaviour are at the core of the competitive balance problematic".

The importance of the concepts has been widely discussed. Fort and Fizel (2004, p. x) claim that "nothing is more important to pro leagues than competitive balance – or the lack thereof". Bourg (2004, p. 4) states that "the notion of uncertainty is the very basis of all competition." Uncertainty of outcome is variously described as "lifeblood" (Dobson and Goddard, 2001, p. 125), "key concept" (Peel and Thomas, 1988, p. 323), and "central concept" (Cairns et al., 1986, p. 5) of professional team sports. Scully (1989, p. 75) claims that "uncertainty of outcome is a necessary feature of competitive team sports". For example, Downward and Dawson (2000) emphasise that in relation to the Bosman verdict, the importance of competitive balance and uncertainty of outcome (in European team sports) was highlighted by the Advocate General (Lenz).[21]

The importance of the concepts drives the academic interest through a number of issues, such as how to measure them and to analyse their importance as demand drivers. However, the empirical evidence on the effects of uncertainty of outcome on spectator demand is mixed (see, for example Szymanski and Kuypers, 1999; Szymanski, 2003b). This may in part be caused by a failure to control fully for the differences in drawing power across teams. Typically, teams from big cities have the biggest fan bases and are therefore able to generate a higher level of revenues, and hence attract the best players (Gerrard, 1999). Further, Gerrard (1999, p. 274) claims that:

[20] However, this is not supported empirically in Forrest et al. (2006), finding no impact from uncertainty of outcome on TV rating.
[21] See also footnote 2 in Szymanski (2001, p. F70).

The dynamics of team competition will force teams to acquire the best possible players and, inevitably, the teams with the bigger resources will dominate, creating dynasties and doormats.

Recognising that the (potential) interest of the different teams are unequal and that demand studies find that sports fans are significantly attracted to winning teams, one can argue that improved sporting success among these teams can be positively related to the total attendance in a league.[22] However, this relationship might only be valid up to some threshold, since certainty of victory can reduce spectator interest (Jennett, 1984).

A reasonable level of competitive balance, preventing individual teams from achieving a level of dominance that would damage the league (Dobson and Goddard, 2001) has also been an important objective for governing bodies. Cross-subsidisation policy is at least partly motivated by improving/maintaining competitive balance. This aim to be pro-competitive differs from other industries (Gerrard, 2000c). The governing bodies have introduced market regulations, which would be deemed restrictive practices and anti-competitive in other industries (Szymanski and Kuypers, 1999). For example, regulations have been introduced to promote joint profits among teams in the league. This is cartel behaviour and usually illegal.[23]

Dobson and Goddard (2001) claim that in the free-market equilibrium, the small-market teams will have negative financial results. The governing bodies will therefore try to protect the league in such a way that it is both sporting and financially viable. This means that it has to promote the survival for these "weak drawing teams" (Sloane, 1971; Fort and Quirk, 1995). However, there are discussions about how regulations affect competitive balance. Cross-subsidisation policy related to competitive balance in professional team sports might also lead to enforcement problems, different incentive effects and conflicts, both within the industry (e.g. maximum wages) and between the industry and the public (e.g. transfer market restrictions and collective sale of broadcasting rights).[24]

---

[22] Survey over variables affecting attendance can be found in for example Cairns et al. (1986), Downward and Dawson (2000), Dobson and Goddard (2001) and Borland and Macdonald (2003). These references have also surveys over attendance studies.
[23] This is described in Balfour and Porter (1991, p. 9): "Recognizing this distinction, professional sports leagues have requested antitrust exemption from collusive restrictions based on alleged business necessity".
[24] See, for example, Quirk and Fort (1992), Scully (1995), Downward and Dawson (2000), Dobson and Goddard (2001) and Szymanski (2001).

Even if Szymanski (2001, footnote p. F69) and Barros et al. (2002) question if the uncertainty of outcome hypothesis might in fact be a modern phenomenon, since: "the Romans, for instance, appear to have enjoyed the unbalanced contest between lions and Christians", there seems, in general, to be at least two main schools regarding the relationship between competitive balance/uncertainty of outcome and fan interest in a league. This is also the view in the review of this literature by Borland and Macdonald (2003). One group claims that there has to be a minimum level of competitive balance, but the general correlation between competitive balance and attendance is limited. Borland and Macdonald (2003) consider this group to argue largely from empirical evidence. The alternative view is that improvement of the competitive balance in general will improve fan interest for the league, through the uncertainty of outcome hypothesis. According to Borland and Macdonald (2003, p. 487), this group has "competitive balance as the 'gold standard'".

On basis of the discussion above, and the concepts by themselves and the question of how to measure them, competitive balance and uncertainty of outcome have been central in the development of economic analyses of team sports (see, for example, Cairns et al., 1986; Bourg, 2004).

2.3. Uncertainty of Outcome

As we have seen, the concepts of competitive balance and uncertainty of outcome are closely related, and many sport economists seem to use them interchangeably. Confusion about these two concepts, including how to distinguish them from each other, is one reason for the high interest in competitive balance in the literature. This thesis argues for having clear definitions of the two concepts to be better able to understand them, and hence achieve more meaningful interpretations of what is meant by them, and their differences and similarities. The conceptual framework of uncertainty of outcome will provide an important basis for the development of the understanding of competitive balance later in this chapter.

It is difficult to find a common definition of uncertainty of outcome in the literature. Downward and Dawson (2000, p. 131) summarize this, by claiming that there is a "difficulty of deciding what precisely is meant by uncertainty of outcome". They also claim that uncertainty of outcome is a frequently used term without a "careful delineation of its precise meaning" (p. 131) and that this is so, even though "researchers are ultimately forced to produce working definitions acceptable to other economists and capable of being observed" (p. 131). The difficulties making a precise definition of the expression have forced Downward and Dawson (2000, p. 131) to expect that: "there is little chance that any writer will produce a definition of uncertainty of outcome that will gain universal acceptance, still less one that wins long-term adherence". This fits with Jennett (1984) and his description of the complexity of uncertainty of outcome.

The expression "uncertainty" indicates that the uncertainty of outcome concept concerns *expectations*, and should therefore (only) be in ex ante form.[25] This interpretation can be found in the more recent European literature on professional team sports, arguing that it, in general, has to do with (un)predictability. Dobson and Goddard (2001, p. 126) use the expression "unpredictability", and Gerrard (2004b, p. 43-4) describes uncertainty of outcome as "the degree of individual contest and tournament predictability". Forrest and Simmons (2002, p. 229) claim that "by uncertainty of outcome is meant a situation where a given contest within a league structure has a degree of unpredictability about the result and, by extension, that the competition as a whole does not have a pre-determined winner at the outset of the competition".

A sporting contest can be related to a number of outcomes. For example, a single contest may be part of an overall tournament. When aiming at creating a definition of uncertainty of outcome, it is necessary to include these aspects. This thesis will therefore suggest a definition that takes into account that it is an ex ante concept related to the probability distribution of different outcomes: *Uncertainty of outcome is the probability distribution for the alternative outcomes of a specified sporting contest.*

This definition therefore takes into account the fact that there is more than one outcome (prize) in a league tournament. In other words, it also includes uncertainty associated with qualification for post-seasonal playoffs and avoiding relegation. An important implication from this thesis is the treatment of prizes in the context of uncertainty of outcome (and later competitive balance). This means that, on the seasonal and long-run

[25] This is also emphasised in for example Owen and Weatherston (2004), who claim that uncertainty of outcome is an ex ante concept.

level, there is usually more than one prize to take into consideration. This is why this thesis suggests expanding the term "outcome" to "outcomes" in the definition.

The importance of separating different dimensions of uncertainty of outcome is, according to Cairns et al. (1986), both related to hypothesis and tests, and as a consequence, this may imply policy decisions.[26] Sloane (1971) is the first to emphasise that uncertainty of outcome in professional team sports consists of more than one dimension, when including long run uncertainty about championship winners (p. 124): "the uncertainty of the length of a team's winning run". In general, the literature has split the term uncertainty of outcome into three main dimensions; match, seasonal and long-run uncertainty of outcome (see, for example, Cairns et al., 1986; Peel and Thomas, 1988; Downward and Dawson, 2000; Dobson and Goddard, 2001; Borland and Macdonald, 2003).[27] Jennett (1984) claims that short-run uncertainty of outcome concerns competitive balance within a season, while long-run uncertainty of outcome has to do with the question of dominance over time. Short-run uncertainty of outcome consists of two forms; the match uncertainty of outcome and seasonal uncertainty of outcome. In other words, the individual contest can be seen as the short-run match uncertainty of outcome, and aggregated to uncertainty of outcome(s) on tournament level, where one level is individual season and the other is the long-run repeated tournaments.

2.3.1. Uncertainty of Outcome at Individual Contests (Match Uncertainty)

Uncertainty of outcome in individual contests can be defined as *the probability distribution for the alternative outcomes for a particular individual sporting contest.* According to Cairns et al. (1986), Quirk and El-Hodiri (1974) express the following mathematical formula for uncertainty of outcome at individual contests with two outcomes (win and loss): $p(1-p)$, where p is the probability for the home team to win. Consequently, the uncertainty is highest when $p = 0.5$. If there is a possibility of drawn matches, the number of outcomes for a team in a match increases to at least three. However, including drawn matches will not change the general conclusions, i.e. the uncertainty of the outcome will increase initially as the probability of a home win rises

[26] Sloane (2006) discusses the importance of including the different dimensions of uncertainty of outcome.
[27] Szymanski (2003b, p. 1155) uses the term "championship uncertainty" as the third dimension.

(and the probability of an away win falls) but at some point uncertainty will start to decline as the possibility of a home win (and an away loss) becomes more and more certain (Cairns et al., 1986).

The uncertainty of a single match outcome is related to probabilities based on the sporting quality of the two teams, home field advantages and other information available, such as injuries prior to the match. These determinants are among the arguments behind the use of betting odds as a measure of this dimension of competitive balance (Peel and Thomas, 1988, 1992, 1996, 1997; Knowles et al., 1992; Kuypers, 1997; Forrest and Simmons, 2002). Measures based on betting odds might be seen as more appropriate than "first generation" measures of match uncertainty, (mainly) based on past performance (for example, the logarithm of the absolute difference in a priori standing) prior to the match, such as in Hart et al. (1975), which is the first contribution to empirically testing match uncertainty (Downward and Dawson, 2000).[28]

However, measures on match uncertainty based on differences in standings do not differentiate between a match between the top two teams contra a match between the teams ranked seventh and eighth in the league, even if these differences might be important for spectator demand and broadcasting rating. This problem was "solved" by Jennett (1984), introducing "championship" and "relegation" significance variables, which will be referred to as *match significance* variables in this thesis.[29] One way to handle this problem in a demand function is to apply two sets of variables, where uncertainty of outcome at contest level is related to the expected output of the match, while the significance of the match relating to tournament outcome(s) will be handled by the match significance variable(s). Combining those variables will probably give a more realistic correlation between the two matches mentioned above and attendance, rather than only using one variable (contest/match uncertainty). It is important in distinguish between the relative and absolute levels of sporting quality. Uncertainty of outcome at contest level should be related to the relative differences between the two teams in a match. Match significance may, in this setting, be reflecting absolute quality.

[28] If applying betting odds, there are a number of potential biases to take into account (see, for example, Forrest and Simmons, 2002; Dobson and Goddard, 2001). The latter reference includes literature review. More discussions about measures of match uncertainty can be found in Dawson and Downward (2005).
[29] This is also considered in Hart et al. (1975), but Jennett (1984) is the first to systematically focus on it.

2.3.2. Uncertainty of Outcomes on Tournament Level (Seasonal Uncertainty of Outcomes)

The uncertainty of outcomes at the tournament level can be defined as *the probability distribution for the alternative tournament outcomes.* This uncertainty does not only mean uncertainty on the championship title, but also to post-seasonal qualification and relegation issues. Uncertainty of outcome therefore directly matters for governing bodies (Sloane, 1971).

The theoretical understanding of the concept of uncertainty of outcome at contest/match level has been relatively unambiguous in the literature, where problems have been related to empirical issues. Uncertainty of outcomes at long-run (repeated tournament) level has not been a major focus of previous research. This topic will be discussed more thoroughly later in this chapter. The main concerns in the literature have been related to uncertainty of outcomes at tournament level, both empirically and theoretically.

In general, uncertainty of outcomes at tournament level has revolved around three issues; the overall sporting difference among the teams, the closeness of competition around the tournament prizes, and match significance for the single team in respect of these prizes.

Rottenberg (1956) and Jones (1969) focus on the general "playing balance" among all teams, and Borland (1987) and Cairns (1987) apply related measures. This is directly connected to the uncertainty of outcome hypothesis and the fact that a higher level of general uncertainty will be achieved, when there is more closeness in sporting quality among all teams. In addition to the above, Davenport (1969) also focussed on the closeness of the championship race, developing a measure that is more heavily weighted around the differences at the top of league than for the other teams, claiming that it is the championship race that is most important for the fans. It is interesting that contributions in the early literature of demand studies (Demmert, 1973; Noll, 1974) have variables for close competition for playoff positions. In other words, spectator demand is not only hypothesised to be related to the fight for winning the league, but also to the prize to qualify for playoffs. Jennett (1984) emphasises that only focusing on the playing strength among the teams in a league might be too narrowly related to uncertainty at tournament level, because it is not capturing enough of the relationship between uncertainty and attendance, since spectators also receive utility from matches that are significant for "the overall championship race" (p. 179).

Broadly speaking, the literature focuses on the (un)predictability of tournament outcomes, mainly the championship, but also playoffs and relegation. Whitney (1988) and Lee (2004) use the term *playoff uncertainty*, while Cairns (1987, p. 260) claims that "seasonal uncertainty refers to the outcome of the championship". This is probably the statement that has been most widely accepted as uncertainty of tournament outcomes. Kuypers (1997) and Szymanski and Kuypers (1999) explicitly include uncertainty of both championship and relegation, Szymanski (2003b, p. 1155) defines seasonal uncertainty as "a close championship race within a season", and Dobson and Goddard (2001) relate it to uncertainty concerning championship, or divisional outcomes, within each season.[30] These arguments are also recognised in the early literature, by Demmert (1973, p. 62), who emphasises that "an essential aspect of the pennant race is that a close race increases uncertainty of outcome by increasing the number of potential winners".

Kuypers (1997) gives an example of how outcome uncertainty was affected in the English Football League through the introduction of a playoff system for promotion. Previously, the top three teams were promoted automatically from the second-tier division (i.e. the Championship) to the top-tier division (i.e. the FA Premier League). The introduction of promotion playoffs limited automatic promotion to the top two teams only with the third promotion place decided by a playoff tournament between the teams finishing third through to sixth. This process increases the uncertainty of tournament outcomes, and is consistent with Fort and Quirk (1995), who state that introducing a championship playoff will increase the uncertainty of which team will win the championship, and Sandy et al. (2004, p. 19) who claim that: "the elaborate playoffs common in many leagues, whereby the top teams play each other for the title or promotion, will help to sustain fan interest".

Uncertainty of outcomes at tournament level, as described above, is considered as the direct form by Cairns (1990), and is described as an uncertainty where "individuals value the identity of the eventual winner of the championship being uncertain" in Cairns

[30] See also, for example, Cairns (1987).

et al. (1986, p. 6).[31] Downward and Dawson (2000, p. 137) use the term "within-season uncertainty", while Borland and Macdonald (2003, p. 482) apply "intra-seasonal uncertainty" and claim that "at any point during a season, a greater degree of evenness in team winning percentages will mean a larger number of teams are in contention for the play-offs."

The second form that has been treated as uncertainty of tournament outcomes is related to prizes and the differences in interest between matches closely related to prizes and other matches. In other words, the interest for a match between two teams of approximately equal sporting quality might be higher if it is a match relevant for avoiding relegation, compared to a match at the middle-of-the-table (with little or no relevance for the top or the bottom of the league) (Sloane, 1971).[32] Hart et al. (1975, p. 18) point out that there are "key matches", such as the fight for avoiding relegation, which should be taken into consideration in a demand function (at match level).

These arguments emphasise that it is not only the "total" quality of the teams in a match that matters. For example, Knowles et al. (1992) and Borland and Macdonald (2003) construct variables measuring total quality. The latter use a measure based on average ranking of the two teams, while Knowles et al. apply a game behind leader measure, anticipating that this measure will give a picture of the relative importance of the match. This is based on a hypothesised positive correlation between sporting quality of the teams and the importance of the matches and is appropriate in closed leagues, but will only be true in an open league when comparing a top match with a middle-of-the-table match. However, this relationship will not satisfy the relationship anticipated in Sloane (1971), that a fight to avoid relegation might be more interesting than a middle-of-the-table match, even if the sporting quality of the latter match probably is significantly higher.

Based on these arguments, one can say that the relationship between uncertainty of outcomes and demand must be wider than just focusing on uncertainty of outcome at match level, and that measures of absolute quality of the matches cannot provide on their own a satisfactory explanation of observed fan behaviour. This means that the relationship between uncertainty and demand also has to include uncertainty about the

[31] See also Cairns (1988, 1990).
[32] See, for example, Noll (2003), who describes the relationship between the fight to avoid relegation and the interest.

prizes of a league and take into account that a match between two teams in the middle of the table, especially at the end of the season, will be irrelevant to the league's prize structure, while a match between two teams fighting to avoid relegation will be very important, when taking into account the prize "not to be relegated" (or to avoid to get the negative prize: relegation), as well as for two teams fighting for the last playoff place.

Jennett's (1984) championship significance and relegation significance measure the level of importance for the specific prize outcome of a match for a given team. These measures are related to both time and league standings during the tournament stage. Jennett focuses on the time dimension and the relatively greater importance matches at the end of the season have on the final standings (given the points already secured).[33] Others have followed Jennett, using match significance as a measure of seasonal uncertainty of outcome. Janssens and Késenne (1987), Cairns (1987), Borland and Lyle (1992), Dobson and Goddard (1992), Wilson and Sim (1995), Kuypers (1997), and Garcia and Rodrigues (2002) have all used similar measures or measures that have taken Jennett into consideration. Other demand studies, which have also used measures that are related to this interpretation of uncertainty of outcome at tournament level, are Borland (1987), Jones and Ferguson (1988), and Baimbridge et al. (1996).

This form of uncertainty of outcome can be seen as indirect and related to single teams' success. The individual spectator is attracted by their team's success. The hypothesis is that the more teams involved in the race for the championship, the higher aggregate attendance (Cairns et al., 1986; Cairns, 1988, 1990; Downward and Dawson, 2000). Downward and Dawson call this form the "seasonal (team-specific) uncertainty of outcome". In other words, match significance in its direct form can be related to absolute quality, but at aggregate level it can be used as indication of the total uncertainty of the "prize race" at a given stage of the league. The more matches that are of higher prize significance, the higher is the level of seasonal uncertainty in the league. This is supported by Borland and Macdonald (2003), who claim that one form of seasonal uncertainty is a relatively higher number of matches being significant

[33] See also Gerrard (2004b). Going through the different approaches to measuring prize significant measures used in the literature mentioned above is beyond the scope of this thesis. Review can be found in for example Dobson and Goddard (2001) and Downward and Dawson (2000). For Jennett's measure, see Jennett (1984) and Cairns et al. (1986).

indicating "greater degree of evenness in a sporting competition" (p. 486).[34] In other words, it is the "aggregate" championship significance that the league is concerned about (Cairns et al., 1986). Hence, the more matches that have higher championship significance, the more uncertain the championship race. The governing bodies might therefore attempt to improve uncertainty of outcome by introducing systems that maximise the number of teams in contention for the championship (Jennett, 1984).

In summary, this thesis will treat uncertainty of outcomes at tournament level as uncertainty of the overall dispersion of sporting success to include uncertainty over which team will win the championship, which teams will qualify for different post-seasonal tournaments including playoffs and European competitions, and the uncertainty of which teams will be promoted and relegated. There is also a relationship between the issues of uncertainty of tournament outcomes. First, the closer the level of playing strength among the teams is, the tighter fights for prizes are anticipated, and hence, the higher level of aggregate match significance can be expected, ceteris paribus. Second, the higher number of significant prizes to compete for, the higher level of aggregate match significance, ceteris paribus, and hence, more uncertainty of tournament outcomes can be anticipated. The focus on different prizes in the literature, concerning uncertainty of tournament outcomes, will be used later in this thesis, with regards to the competitive balance and competitive intensity concepts.

2.3.3. Uncertainty of Outcome in Long-Run Repeated Tournaments (Long-Run Uncertainty)

The third dimension of uncertainty of outcome is the uncertainty of long-run repeated tournaments (long-run uncertainty) that takes into account the dynamic relationship in repeated tournaments.[35] Typically, a sports league is dynamic in the sense that it is repeated season after season, but examples can also be found in individual sports, such as Tour de France and World Cup in biathlon. The dynamics may be very strong particularly in team sports operated as merit hierarchy systems with promotion and

[34] Borland and Macdonald (2003, p. 486) use the comparison based on "natural experiment" between tournaments the Football/Premier League and FA cup in English football by Szymanski (2001) as an example of a third approach of seasonal uncertainty.
[35] Borland and Macdonald (2003) use the term *inter-seasonal uncertainty of outcome*.

relegation such as European football. In these systems participation in the current season is a dependent on performance in the previous season.

Discussion of long-run uncertainty tends to concentrate on the domination of a league by a small number of teams. Sloane (1971) considers the domination of the Scottish league by Rangers and Celtic and focuses on the negative effect that long-run domination might have on attendance (especially for dominated teams). Sloane's approach has been the norm for understanding uncertainty of outcome in long-run repeated tournaments. This can be seen in various descriptions and definitions of long-run uncertainty of outcome, as relating to domination by one team (Cairns et al, 1986; Cairns, 1987; Peel and Thomas, 1988; Szymanski and Kuypers, 1999; Downward and Dawson, 2000), by one or a few (two) teams (Jennett, 1984; Kuypers, 1997; Szymanski, 2003b; Sandy et al., 2004), and by groups of teams (Borland, 1987; Borland and Lye, 1992).[36] Long-run domination is related to the uncertainty at the beginning of the season as to which team will win the championship based on outcomes in previous seasons (see, for example, Dobson and Goddard, 2001, p. 42).[37] This uncertainty will be low if one team has been the dominant over many seasons.

Given a certain level of competitive balance, the possibility of domination will depend in part on the structure of the league. If winning the championship over a number of years is the criterion for defining a dominant team (for example, Rosenborg BK winning the Norwegian championship thirteen times in a row 1992 - 2004, and Skonto Riga in Latvia winning fourteen in a row 1991 - 2004),[38] a change in the structure of how the league champion is determined may alter a team's probability of winning the championship. One method of reducing the likelihood of long-run domination (i.e. dynasties and doormats) is the introduction of a championship playoff system (Fort and Quirk, 1995).[39] However, long-run repeated participation in the playoffs by the same

---

[36] A number of these examples have assumed that long-run domination is equivalent with long-run uncertainty of outcome.
[37] This is shown very clearly in the description of Szymanski (2003b, p. 1155), who calls this dimension "*championship uncertainty*".
[38] See, for example, www.rsssf.com.
[39] Another effect of such a system is that the uncertainty according to the championship, and hence uncertainty of outcomes at tournament/seasonal level, increases. An interesting example here can be found in the Norwegian top league for women's handball. The dominating team won all matches in the league, as well as becoming the cup winner. However, the team lost one match in the play-off, and it therefore did not qualify for the play-off final (www.vg.no/sport/haandball/03_04/kvinner/ and www.handball.no/p1.asp?p=17020). Consequently, it did not qualify for the EHF Champions League. The opinion found this to be unfair (pub.tv2.no/nettavisen/sport/handball/article221137.ece), as it "conflicts with our sense of justice", in Sanderson and Siegfried (2003, p. 260).

group of teams also represents a form of long-run domination (Borland and Macdonald, 2003).

The meaning of long-run uncertainty of outcome is difficult to define precisely. Is it only concerned with domination by one or two teams at the top of the league? Or should it also explicitly include the performance of other teams in the league? The existence of dominant teams necessarily implies the existence of dominated teams. Concern with the performance of both dominant and dominated teams relates to the uncertainty of performance persistence. The nature of the uncertainty of persistence may change over time within different segments of the league. For example, the nature of the domination at the top of English football has varied over the last 15 seasons. Initially the title race in the FA Premier League was dominated by one team (Manchester United). This domination has gradually expanded to two teams (Manchester United and Arsenal) and then to four teams (Manchester United, Arsenal, Chelsea and Liverpool). Thus there is low uncertainty of outcome as regards which teams are potential winners of the Premiership title at the start of each season but high uncertainty of outcome over which one of the "Big Four" will emerge as league champions.

2.4. Competitive Balance

Chapter 1 characterised competitive balance as complex, and the treatment of uncertainty of outcome in this chapter has further emphasised that real-world leagues are complex structures. One of the aims of this thesis is to create a conceptual framework for the understanding of competitive balance. A simple league context will be used as the basis for the analysis in the rest of this chapter. The arguments will be extended to a complex league structure in the next chapter. There are three reasons for distinguishing between simple and complex league systems. First, it is important to have a common basis for analysis given the huge variety of league systems, current and historic, between and within sports, as well as national and continental differences. There are very significant differences between the organisation of the NAML and that of European football, and, in turn, there are significant differences in the structure of domestic football leagues across the members of the UEFA. Second, it will simplify discussion and understanding of the large number of measures of competitive balance and uncertainty of outcome in different leagues. Third, most of the relevant aspects of the peculiar economics of professional team sports are valid in simple league systems and hence can be analysed more effectively by abstracting from the complexities of real-world leagues that often serve to obscure the key issues.

2.4.1. A Simple League Context

By a simple league context is meant a theoretical league constructed to be as simple as possible. The simple league context provides a common basis for discussion, and allows the analysis to be focused on the core aspects of concept competitive balance free from the confusions created by the complexities of real-world league structures. The simple league context has five defining characteristics:

1. The league is *closed* with no relegation and promotion. It is also assumed that the same set of teams will compete in the league season after season.
2. The structure of the league is *unitary*. This means that the results from the matches are equally weighted to determine the tournament outcome. The tournament consists only of a regular season with no post-seasonal playoffs.
3. A simple league structure consists of a *round-robin match schedule* with all teams in the league playing each other the same number of times home and away. There is no elimination (i.e. knock-out) phase in the league tournament.
4. Individual match outcomes are limited to wins and losses with *no tied matches*. In addition there is no differentiation between whether the outcome is decided in regular time or via some form of overtime tie-breaker. And there is no differentiation based on the margin victory. All wins are awarded the same points with all losses awarded zero points.
5. A simple league has only *one prize*, the championship, where the winner is decided on basis of the best win-loss record over all matches played.

Although all modern league structures are much more complex, the simple league context is representative of the structure of "first generation leagues" from which modern leagues have evolved.

2.4.2. Conceptual Framework

Competitive balance is a complex concept consisting of three principal dimensions; win dispersion, performance persistence and prize concentration.[40] Win dispersion relates to the distribution of wins within single-season tournaments. Performance persistence concerns the relationship between sporting outcomes across seasons. The third dimension is prize concentration, which in the simple league context refers to the concentration of championship winners over time. Whereas win dispersion is concerned with within-season performance, win persistence and prize concentration are both time-dependent, related to performance across seasons.

Competitive balance across seasons depends on the correlation over time, not only for the top teams, but for the whole league (see, for example, Daly and Moore, 1981; Balfour and Porter, 1991; Butler, 1995; Eckard, 1998, 2001; Humphreys, 2002; Szymanski and Smith, 2002). As emphasised in Gerrard (1998), competitive balance depends not only on the distribution of performance of teams within a season but also includes the degree of stability in the performance for teams across seasons. A league with a high level of competitive balance should not have the same teams performing successfully season after season (i.e. dynasties) and the same teams performing weakly over many seasons (i.e. doormats). Lower competitive balance over time can therefore be reflected in less variation in the teams' performance over time, with all teams having a high probability of repeating their current performance levels in subsequent seasons (Eckard, 1998, 2001). In such a league, the variation of the individual teams' winning percentages may be tend to be lower but not necessarily so. The league may also be characterised by strategic groups (Caves and Porter, 1977) consisting of groups of teams with similar levels of performance that persist across seasons.

In addition to these three basic dimensions, the literature also tends to highlight domination by teams from large markets as a separate aspect of competitive balance (see, for example, Vrooman, 1996, 2000). However, the view adopted in this thesis is that domination by large-market teams is a time-dependent phenomenon captured by win persistence and prize concentration.

[40] See La Croix and Kawaura (1999). There are also examples in the literature that the term competitive balance has been used on match level. This point will (only) be included in the discussion of uncertainty of outcome at match level, and is hence not treated as a dimension of competitive balance in this thesis.

Scully (1989) claims that measuring the distribution of championships is limited, since such measures are binary – either a team wins the championship or it does not – and thus cannot capture the closeness of the championship contest. This is also emphasised in Szymanski and Kuypers (1999), who claim that it is possible for a league competition to be extremely close and exciting each year and yet still be won by the same team. On the other hand, it is also possible that a number of teams could win the league without any of the championship races being close and exciting. In this case fan interest might be high at the start of the season with the anticipation of several credible contenders for the championship, but the interest may dissipate through the season as the championship contest becomes a "one-horse race". However, by recognising the potential importance of the three different dimensions, a fuller understanding of a particular league's competitive balance can be gained, thus avoiding the problems highlighted by Scully and Szymanski and Kuypers that result from focusing attention on only one specific aspect of competitive balance. Hence championship concentration can give valuable insights, particularly when combined with measures of the other dimensions of competitive balance.

The complexity of competitive balance can be exemplified by looking at the Norwegian league, where the same team, Rosenborg BK, won 13 championships in a row from 1992 to 2004. It is easy to conclude that this league must have weak competitive balance because of the long-run domination. However, it does not say anything about the closeness of each championship, neither the number of teams competing every season for the championship title, nor the "turnover" in the identity of the teams finishing second and third. For example, when Rosenborg won the 13th title in a row, the final standing shows that the championship was won by only one goal. In addition, the distribution of sporting performance, when taking all teams in the top division into account, was well distributed, measured by the NSQF ratio (see below for further explanation of the source and definition of the NSQF ratio).[41]

According to Buzzacchi et al. (2003), the literature of competitive balance has mostly been concerned with static measures[42] of the equality of winning opportunities for

[41] The calculated index was 0.99. Quirk and Fort (1992) indicate that an index of unity shows perfect competitive balance, and the higher value of the index, the weaker competitive balance.

[42] This term is used by Szymanski and Smith (2002) as a distinction from dynamic competitive balance. Koning (2000) uses it as a description related to within-season competitive balance.

teams in either individual matches or the overall championship. Buzzacchi et al. assume that fans care about competitive balance in the sense that they want winners to change from time to time (although they may also care about the variance of success among the teams within the season). More teams in the championship contest increase fan interest among followers of those teams, as well as "team-neutral" fans, compared to a league without a close contest for the championship. Buzzacchi et al. (p. 174) claim that "if each team experiences diminishing returns to success in terms of fan interest, then a league with greater competitive balance in this dynamic sense will be more successful." These considerations are also taken into account later in this thesis.

The general definition of competitive balance can be expanded to include both ex ante and ex post time considerations. In other words, competitive balance will be treated as a wider concept than (the ex ante) uncertainty of outcome. Some of the newer literature has explicitly considered the concept of competitive balance to be related to expectations of the future. Gerrard (2004, p. 44) suggests that "competitive balance refers to the shape of the probability distribution of the likelihood of teams winning the tournament." Szymanski (2001) and Buzzacchi et al. (2003, p. 168) claim that competitive balance is concerned with the *expectation* of "who will be the winner". Putting the ex ante (time) dimension into the general definition of competitive balance yields the following definition: *Ex ante competitive balance is the probability distribution of sporting outcomes between the teams competing in a league/tournament.*

Ex ante competitive balance is related to expectations based on currently available information, such as the strength of the teams in a league and the schedule system. Often the information is qualitative and imprecise. For example, Gerrard (2004b, p. 45) claims that "little evidence exists on the distribution of playing talent because of the inherent difficulties in measuring it." In theoretical terms, competitive balance has been treated as both an ex ante and ex post concept, whereas empirical measures have focused almost exclusively on ex post outcomes. This is emphasised in Gerrard (2004b), claiming that measures of competitive balance are mainly based on either the ex post distribution of sporting outcomes or the distribution of the economic size of the teams. However, there is a connection between ex ante and ex post competitive balance, since the information from current and past performance provides data for determining expectations of the next season's (ex ante) competitive balance. The long-run uncertainty of outcomes concept can be related to ex ante across-seasonal competitive balance. Turning it around, ex ante competitive balance in relation to performance persistence and prize concentration can be seen as an element of the uncertainty of outcome in long-run repeated tournaments. A similar relationship is to be found between ex ante win dispersion competitive balance and the uncertainty of outcome at tournament level regarding overall dispersion of sporting outcome, as is also emphasised in Fort (2007, p. 648) referring to Dawson and Downward (2005).

Indeed, there are examples in the literature of ex post definitions of competitive balance,[43] such as the distribution of performance (Hall et al., 2002), the distribution of team wins (Depken, 1999; Larsen et al., 2006), the distribution of winning percentages (Marburger, 2002; Maxcy, 2002), and the distribution of league championships (Fort, 2003). In other words, the definition of ex post competitive balance is based on the distribution of actual outcomes in a league/tournament and hence is closely associated with the empirical measures of competitive balance.[44] *Ex post competitive balance* is defined in this thesis as *the distribution of actual sporting outcomes between the teams in a league/tournament.*

A general overview on the understanding and suggestions of the concept of competitive balance, taken from the discussions above, is shown in Figure 2.1.[45]

[43] Again it should be emphasised that the descriptions related to the concept competitive balance are mine.
[44] The descriptions are collected from literature that uses them explicitly to describe competitive balance.
[45] This figure takes into account many of the points related to complexity of the concept competitive balance in Gerrard (2006a).

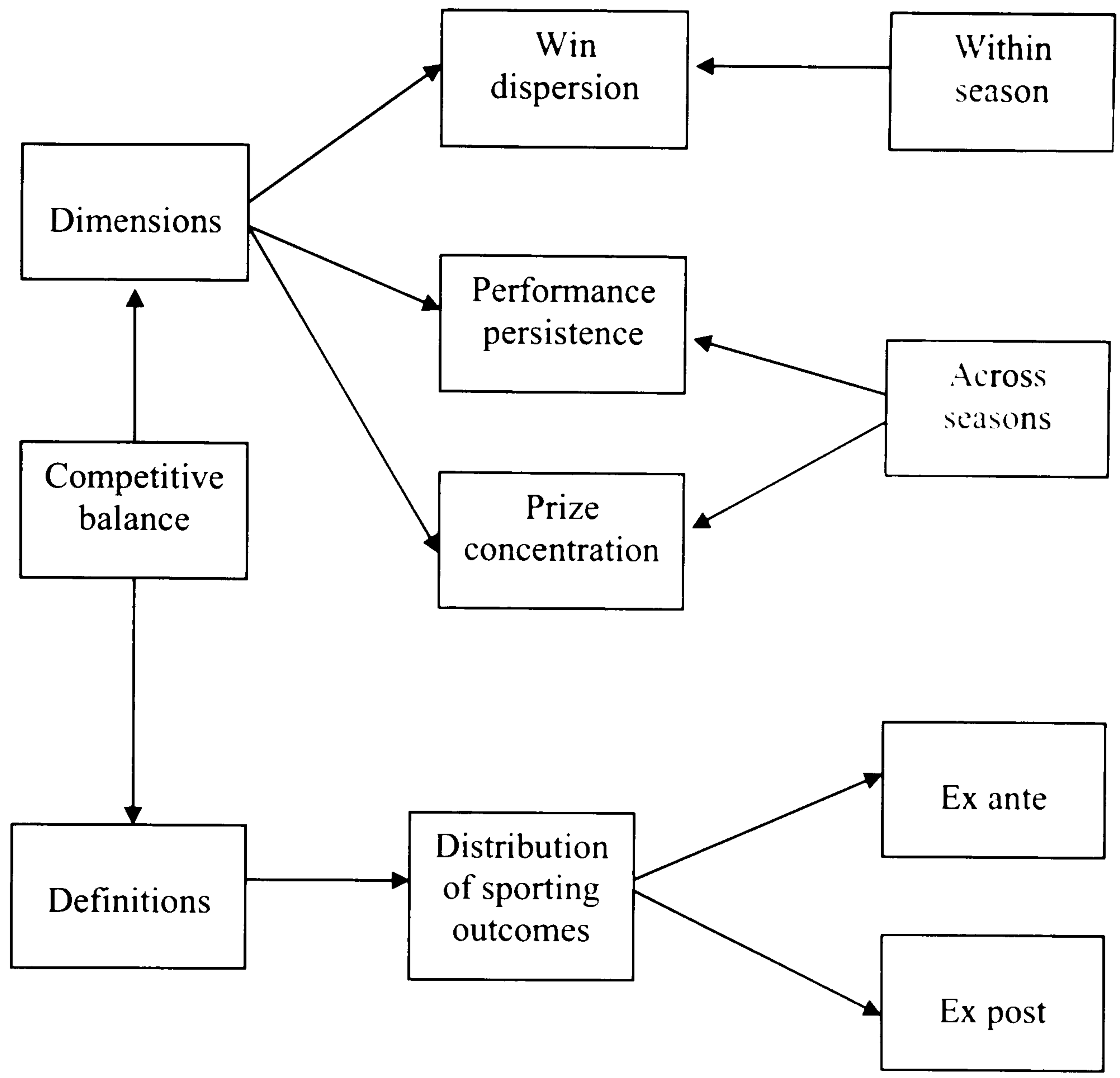

Figure 2.1: Overview of competitive balance

The definition of competitive balance by the Blue Ribbon panel on Baseball Economics (Levin et al., 2000, p. 5) has been categorised as "qualitative", by for example Zimbalist (2002, p. 113) and Sanderson and Siegfried (2003, p. 257):

> In the context of baseball, proper competitive balance would be understood to exist when there are no clubs chronically weak because of MLB's financial structural features. Proper competitive balance will not exist until every well-run club has a ***regularly recurring hope of reaching postseason play***.

It is difficult to use this description as a definition of the concept of competitive balance, because it seems to be more related to a proposed minimum level of what is considered to be an appropriate level of competitive balance in a league. There are varying views of this "definition" in the literature. Szymanski (2003a) claims that this description of competitive balance "is not easily measurable" (p. 471). Humphreys (2003) calls the definition a "temporal dimension of competitive balance" (p. 285), while Kahane (2003) describes its indistinctness as follows (p. 289):

> If a so-called well-run team reaches the post season at least once in every 5 years, is this sufficient for competitive balance? If it is, then the alleged competitive balance problem in MLB can be solved quite easily by following the National Hockey League's paradigm by increasing the number of teams eligible for postseason play.

Further Kahane (2003, p. 289) argues that:

> The bottom line is that there are many notions of competitiveness including a team's ability to win a championship, some given proportion of its contests, a divisional title, any individual contest, or to improve its chances for any of the former from season to season. Without a clear definition of competitive balance, the discussion of whether leagues are sufficiently balanced has little meaning.

It is exactly these concerns of Kahane that are taken into account in this thesis. The many notions of competitiveness are summarised in the term *competitive intensity* (see Chapter 3), as well as in the analysis of competitive balance in a complex league context (see Chapter 3). Providing a clear and comprehensive meaning of the concept of competitive balance is the central objective of this thesis. The definition of competitive balance used by the Blue Ribbon panel on Baseball Economics and the arguments by Kahane will be followed up in Chapter 3.

Several researchers have described competitive balance relative to some ideal or extreme case such as perfect competitive balance, perfect competitive dominance and optimal level of competitive balance. The cases are of course only theoretical conditions. By contrast, the Blue Ribbon panel (Levin et al., 2000) has given a definition of competitive balance more in terms of a minimum acceptable level of competitive balance. Similarly, Jones (1969) suggests an "equilibrium" solution where competitive balance is sufficient to ensure the viability of the league. "Although it will not be one which results either in joint profit maximization for the group or attempted profit maximization for each club, it will ensure group "equilibrium"" (Jones, 1969, p. 4).

This thesis focuses on three dimensions of competitive balance and implies that an improvement in competitive balance (in a simple league context) is associated with a narrower distribution of win dispersion, weaker performance persistence from season to season, and less prize concentration. The polar cases, perfect competitive balance and perfect competitive dominance, will be treated below, while discussion of the optimal level of competitive balance is beyond the scope of this thesis.

For the win dispersion dimension, perfect competitive balance implies that every team has an equal probability of winning each match, so that every team is expected to win half of their matches ratio: $E(w_i) = 0.5$, where w_i is the win ratio for team i, $i = 1, \ldots, N$, and N is the total number of teams in the league.

Perfect competitive balance in the performance persistence dimension implies zero correlation between the distribution of outcomes in season t and season t – b:

$$E[\text{corr}(CB_t, CB_{t-b})] = 0, \; b > 0$$

where CB_t is the distribution of match outcomes in season t.[46]

Prize concentration in perfect competitive balance in a simple league context implies an equal distribution of championship titles between all competing teams: $E(C_i) = T/N$, where C_i is the number of championships won by team i, T is the number of seasons and N is the total number of teams in the league. This is consistent with Forrest and Simmons (2002), who use an equal number of championship wins by all teams in the league, during a specified period of time, as an indication of a perfectly balanced contest, implying that the frequency with which a team wins the championship is once every N seasons (Demmert, 1973).[47] The more closely the actual distribution of championship titles for each team is concentrated around the ideal of T/N, the better is a league's competitive balance, ceteris paribus.

[46] This is consistent with Groot (2008), arguing that perfect competitive balance is reflected by zero correlation, and not perfect negative correlation, where the same team would be the champion every second season.
[47] See also Rottenberg (1956).

It should be noted that in the case of perfect competitive balance, there is no time dependency in outcomes. In terms of the uncertainty of outcome, this means that the probability of winning the next match is independent of all previous matches, and that the probability of winning any particular match in one season is independent of results in previous seasons. In other words, if there is perfect competitive balance in a league, each team has the same probability of winning a specific match in the final stage of the tournament as at the start of the season. The same holds for the probability of winning the championship overall. These relationships are also emphasised in Borland and Macdonald, referring to Leifer (2000, p. 11):

> Competitive balance yield winners and losers in both games and seasons, but it keeps open the chance that winners will lose and losers will win in subsequent competition. This helps undermine the significance of past winning and losing by arousing public interest in upcoming competition, no matter what has happened in the past.

Forrest and Simmons (2002) look at alternatives of outcomes in a perfectly balanced contest. For the win dispersion dimension, they consider the case of every match ending as a tied contest[48] (or the same outcome/equal score - Bourg, 2004) as well as the case of each team winning and losing half of their matches (see also Horowitz, 1997; Demmert, 1973). These suggestions can also be found in Rottenberg (1956). In the simple league context the case of tied matches is excluded. This is dealt with in the discussion of more complex league structures. The relationship between home field advantages and competitive balance also needs to be considered. An equal number of wins and losses for all teams, with teams winning all of their home matches, gives perfect competitive balance, but this negates the uncertainty of outcome hypothesis, since although there is an equality of win ratios between teams, there is no uncertainty of match outcome for individual matches.[49]

Sloane (1976a, 1976b) emphasises that the win ratios of teams should be randomly distributed around the mean. This is allowed for in the notion of an idealized standard deviation of win ratios based on the normal distribution in the NSQF ratio, the most

[48] Given a complex league context.
[49] This is based on an answer from Prof Dr Bill Gerrard in a discussion among the audience after a presentation about home field advantages at the NASSM 2007 conference.

widely used measure for within-season competitive balance (see below for further explanation of the source and definition of the NSQF ratio). When describing perfect competitive balance, one should therefore remember the differences between the ex ante and ex post time dimensions, since it is at the ex ante level equality exists in regards to equal probabilities of outcomes and not equality of the ex post actual distribution. Perfect (ex ante) competitive balance is related to equal outcome probabilities. This means that equality of observed (ex post) outcomes is only expected when the number of repetitions tends towards infinity. Given the simple league context, the observed outcome should have a binominal distribution on basis of equal probability for success and failure in any individual event. This should be reflected in the statistical properties for all three dimensions of competitive balance.

By analysing the ex ante and ex post aspects of perfect competitive balance, the potential problem that "It would be impossible to establish a ranking, and the interest of such a championship would be debatable", described in Bourg (2004, p. 5) for a league containing perfect competitive balance might *not* be a relevant problem for professional sports leagues, because subjective expectations (probability distributions) by individuals will create (subjective) rankings.

Perfect competitive dominance occurs when every team is certain (i.e. unit probability) of beating every weaker team in every season. For a single season, this would imply that the strongest team wins all its matches, the next strongest team wins all of its matches apart from the matches against the strongest team, and so on. The weakest team loses all of its matches (see, for example, Utt and Fort, 2002 and their references). In respect to performance persistence, the same pattern of outcomes is repeated season after season. This means that, when in the case of perfect competitive dominance, the strongest team has a unit probability of winning the championship every season (Szymanski, 2001).

2.5. Measures of Competitive Balance in a Simple League Context

Empirical analyses of competitive balance have always been a relevant part of the literature. The studies are related to both general overviews of changes in competitive balance over time in single leagues, as well as comparisons to other leagues (see, for example, Quirk and Fort, 1992; Gerrard, 2004b). A number of studies have analysed the effects on competitive balance of changes in institutional rules (see, for example, Fort and Quirk, 1995) and the relationship between competitive balance and attendance (see, for example, Schmidt and Berri, 2001).

A key topic in the literature is how to measure competitive balance. This concern is of course driven by the importance of quantifying competitive balance for empirical analysis, but also by the difficulties of finding adequate measures, in part due to the complexity and conceptual difficulties described in the previous sections. As a consequence of the multi-dimensional complexity, it seems impossible to construct one perfect measure of competitive balance. However, there still seem to be many sport economists trying to find the perfect measure of competitive balance. This thesis adopts the position that there is no perfect single measure of competitive balance. In this section, the wide range of measures of competitive balance will be treated in the context of a simple league system using the three dimensions of win dispersion, performance persistence and prize concentration. Using a multi-dimensional approach recognises the warning by Szymanski (2003b) that many measures, such as the seasonal standard deviation, do not capture the dynamics of competitive balance over time. In addition to the one-dimensional measures presented below, Eckard (1998, 2001) and Humphreys (2002) have constructed measures combining two dimensions - win dispersion and performance persistence (see also Eckard, 2003). These measures combine "dynamic" changes over time and the "static" within-season win dispersion in a single measure.

2.5.1. Win Dispersion

Measures of win dispersion can be seen as measures of the second moment of a distribution. Several are based on the dispersion of sporting outcomes among the teams in a league. Such measures have been used both for competitive balance and uncertainty of outcome at the tournament level (i.e. seasonal uncertainty of outcome).[50] The basis for these measures is the dispersion of outcome, as suggested by Rottenberg (1956).[51] Generally, competitive balance improves if the win dispersion narrows.

[50] Since uncertainty of outcome and competitive balance have some common points, a number of the measures presented as competitive balance measures can originally have been constructed in the context of uncertainty of outcome.
[51] Rottenberg (1956, p. 246): "The dispersion of percentages of games won by the teams in the league".

One group of win dispersion measures is related to **statistical measures of win ratio dispersion**. The *range* is the simplest, and takes only into account the two extremes - the best and the weakest teams' performance - and ignores performance by the intermediate teams (Quirk and Fort, 1992). This measure is also used by Borland (1987) and Lee (2004). An alternative measure is the *interquartile range*.[52] This is the difference between the first and the third quartile, but again only part of the distribution of outcome information is taken into account.

The next measures of win dispersion are based on *dispersion of outcome (standard deviation/variance)* for all teams in the league. Rottenberg (1956) and Jones (1969) obviously focus on the general distribution of outcome among all teams. The first researchers to explicitly use of the variance of performance measure were Borland (1987) and Cairns (1987).[53] Compared to the range, a general advantage with these measures is that they reflect the competitive situation for all teams. The formal definition of variance in the case of win dispersion is:

$$\sigma^2 = \frac{\sum_{i=1}^{N}(w_i - 0.5)^2}{N-1}$$

where w_i is the win ratio for team i, and N is the total number of teams in the league. In the case of the simple league context, the average win ratio ($\overline{w}$) is 0.5. There might be a discussion whether the denominator should include a reduction in the degree of freedom or not. In the purpose of summary statistics, reduction of the denominator is not necessary, because the average win ratio in the simple league context is known. However, in the context of calculating a level of competitive balance, the purpose is inferential, so it can be used to compare competitive balance either across seasons in own league and/or other leagues. Sporting contests and tournaments are stochastic because they involve random and unexplained variation. It is, hence, conventional to use N-1, as is also consistent with the literature.

---

[52] It can also be called semi-interquartile range, as in Copeland et al. (2005).

[53] Demmert (1973) also considers the variance on winning percentages among the teams. However, this measure had a different purpose than win dispersion, as it was aimed to be a measure for the intensity of pennant race in a demand study for the MLB. Because it considered the weaker teams (out of the race) too much, it was omitted from the final demand equations.

Usually, the dispersion is reported in the terms of standard deviation, such as in Scully (1989):[54]

$$\sigma = \sqrt{\frac{\sum_{i=1}^{N}(w_i - 0,5)^2}{N-1}}$$

Even in the simple league context, there is a problem in comparing competitive balance between leagues with different numbers of teams. Scully (1989, p. 89) claims that different numbers of matches between seasons might affect the "expected" standard deviation. "The fewer number of games played between contestants, the more likely is the impact of random factors in the outcomes." In other words, measures based on summary statistics might be difficult to apply when comparing tournaments of different numbers of teams and matches, since many of these measures are dependent both on unit and number of observations. Therefore, it is suggested that one standardises summary statistics in relation to a given benchmark. In the context of competitive balance, this has usually meant calculating the ratio of actual (absolute) measure of competitive balance relative to the value of that measure in the ideal case of perfect competitive balance.

This is what is done in the most popular measure of competitive balance, based on end-of-season outcomes, and designated in this thesis as the *NSQF ratio*.[55] The suggested name of the measure is related both to Quirk and Fort (1992) as the first to apply the measure, but also to Noll (1988) and Scully (1989), since the measures is, as emphasised in Quirk and Fort, "suggested by Noll (1988) and applied by Scully (1989)" (p. 244). However, a number of names have been used in the literature. Fort (2007) makes it clear that the measure is developed by Noll and Scully, and emphasises that it therefore should not only be referring to Quirk and Fort (1992). Even if Fort (2007) follows Lee and Fort (2005) applying the name RSD (ratio of standard deviation), this thesis prefers using the names behind the measure, because this will reduce the probability of confusion. This competitive balance measure is constructed on basis of an idealized standard deviation, calculated using perfect ex ante competitive balance as the

[54] For transformation of the standard deviation measure into a zero to one scale, where perfect dominance is reflected by zero and perfect competitive balance is equal to unit, see Groot (2008).
[55] This measure is the most popular indicator of competitive balance in the North American literature (e.g. Szymanski, 2001; Dobson and Goddard, 2001).

comparative benchmark. This measure takes account of differences between leagues as regards the number of teams and matches. The NSQF ratio is defined as follows:

$$\text{NSQF ratio} = \frac{\text{Actual standard deviation}}{\text{Idealized standard deviation}} = \frac{\text{ASD}}{\text{ISD}} = \frac{\sqrt{\dfrac{\sum_{i=1}^{N}(w_i - 0.5)^2}{N-1}}}{\dfrac{0.5}{\sqrt{m}}}$$

where m is the number of league matches per team during the season.[56] Expressing the actual standard deviation (ASD) as a ratio of the idealized standard deviation (ISD) provides a basis for comparing the degree of competitive balance in different leagues.[57]

Goossens (2006) uses a different type of benchmark, when constructing the *NAMSI (National Measure of Seasonal Imbalance)* as her proposed measure of competitive balance. Rather than applying the idealized competitive balance (as a measure of perfect competitive balance), she uses the opposite polar situation of perfect competitive dominance as the benchmark.

Another group of win dispersion measures uses only **part of the win ratio distribution** and compares the actual dispersion of a specified portion of the win ratio distribution with the theoretical idealised dispersion for that portion. The most common form of this type of measure is the *excess tail frequency* of the win ratio distribution (Quirk and Fort, 1992; Fort and Quirk, 1995; Lee, 2004). Excess tail frequency is the difference in the actual number of observations in the tail of the win ratio dispersion and the idealised number of observations that would be expected lie in that tail under conditions of perfect competitive balance. Lee and Fort (2005) and Fort and Lee (2007) apply this measure in a logarithmic variant, called LTL, *the Log of the Tail Likelihood.* Weaker competitive balance may be reflected in excess tail frequencies both at the top of the league and the bottom of the league.

Gerrard (1998) constructs a *concentration ratio*,[58] comparing the actual k-team concentration ratio for top k teams with the idealized concentration ratio for the same number of teams if every team had a win ratio of 0.5 (in a simple league context).

---

[56] Downward and Dawson (2000, p. 63) show how to measure the idealized standard deviation using a binomial model.

[57] The short forms ASD and ISD are used in Fort (2007).

Based on a five-firm concentration ratio from the "standard industry", Michie and Oughton (2004) apply a *five-club concentration ratio (C5 Ratio)* in European football. In a simple league context, this equates to the CR(5) used by Gerrard (1998). Adjusted for the number of teams in the league (division), Michie and Oughton extend the formula to give the *Index of Competitive Balance (C5ICB)*:

$$C5ICB = \frac{C5\ Ratio}{5/N} * 100$$

where the C5 Ratio, in the simple league case, can be defined as the share of wins by the top five teams as a ratio of the total number of wins in the league.

Koning (2000) measures the concentration ratio by comparing the actual performance by the top k teams with the theoretical best attainable performance by these k teams (and similar to the notion of perfect competitive dominance used in Goossens, 2006).[59] Also Goossens and Késenne (2007) measure concentration of the k better teams (big market teams). They measure the difference in aggregated win percent for the top k (large) teams and their theoretical minimum win percent, as a share of the difference between their theoretical aggregated maximum and minimum win percent.

Alternatively, the concentration ratio can be reported as a ratio for the weakest teams, where relatively small concentration of wins among these teams can reflect weaker competitive balance in the league.

Win dispersion can also be measured on basis of measures focusing on the level of **inequality between the observed win ratios**. Depken (1999, p. 207) refers to a number of industrial studies measuring concentration and relative "competitiveness", citing Kamerschen and Lam (1975) as a general overview of such measures. Depken (1999, 2002) defines the *Herfindahl-Hirschman Index (HHI)* as:

$$HHI = \sum_{i=1}^{N} (MS_i)^2$$

[58] The model is originally based on a point score model.
[59] The model is originally based on a point score model.

where MS_i is the market share of the i^{th} firm on a scale of 0 to 1.

In the context of professional team sports and competitive balance, MS_i can be treated as team i's number of wins as a share of total number of wins in the league (Depken, 1999; Downward and Dawson, 2000). As emphasised in these references, HHI is sensitive to the number of teams in the league. Perfect HHI is 1/N, where N is the total number of teams in the league. Depken (1999) uses actual HHI minus idealized HHI (see also Eckard, 2001), and defines:

$$\text{dHHI} = \text{HHI} - 1/\text{N} = \sum_{i=1}^{N}\left(\frac{2 * wins_i}{N * m}\right)^2 - \frac{1}{N}$$

Downward and Dawson (2000) measure a "modified Herfindahl Index (MHI)" as:

$$\text{MHI} = \text{N} * \text{HHI}$$

A similar measure is used by Michie and Oughton (2004) for leagues with a point score system. Converting the measure to a simple league context is done by multiplying the MHI by 100.

Instead of using perfect competitive balance as the benchmark, Gerrard (2004b) uses perfect competitive domination. Gerrard measures the Herfindahl index on the concentration of the top four teams, adjusted for maximum possible Herfindahl index, which is 0.25, if the same teams have top four finishes every season.

Entropy is a measure used for uncertainty in information theory, and is applied in the context of competitive balance by Horowitz (1997), measuring relative entropy as the actual (absolute) level of entropy standardised by maximum entropy. *Gini coefficients* (and *Lorenz-curves*[60]) are widely used as "conventional economic measure of inequality" (Schmidt and Berri, 2001, p. 147), for example as indicators of the distribution of wages. They are also relevant in the context of professional team sports. The Gini coefficient, for example, is used by Gerrard (1998), Schmidt (2001) and Schmidt and Berri (2001) as a measure of within season competitive balance. The important clarification in Utt and Fort (2002) about the peculiarity of professional team

[60] See, for example, Michie and Oughton (2004).

sports should be noted when it comes to measures of perfect competitive domination on seasonal level. Since any individual team can only win its own matches and hence cannot win every match played in the league, perfect competitive dominance cannot be treated in the same way as pure monopoly and measured as a unit "market share" of wins. The same holds for the HHI measure (Depken, 1999).[61] Utt and Fort (2002) point out that Fort and Quirk (1997) avoided this problem by benchmarking against perfect competitive dominance in the win dispersion dimension. This *adjusted Gini coefficient* is illustrated in Utt and Fort (2002, p. 370).[62] It is also applied in analyses of the NFL by Larsen et al. (2006).

By ranking each team in a league, Haugen (2008) measures the aggregated squared differences between the theoretical score for each team in a perfect competitive dominance situation and their actual score, as a percent of the aggregated squared score differences between perfect competitive dominance and perfect competitive balance for this league.[63]

Different variants of *games behind the leader (winner)* measures are common among the early contributions in the field of the economics of professional team sports (Davenport, 1969; Demmert, 1973; Noll, 1974; Balfour and Porter, 1991; Ross and Lucke, 1997). Instead of using games behind the leader (winner), the degree of competitive balance can be indicated by simply focussing on measures of the *performance of the best team* (see, for example, Gerrard, 2004b; Groot, 2008).

A recent contribution by Groot (2008) is the *surprise index*, measuring the realized surprise points as a ratio of the maximum number of surprise points. The surprise points are measured on basis of the results within the season, where the level of surprise increases with the distance in the end-of-season rank, when a weaker team wins against a stronger team.[64]

The measures above are all ex post measures (that might be used as a basis for the ex ante measure for season t+1). There are no published examples of measures of *ex ante*

---

[61] See also Larsen et al. (2006) for discussions of HHI in the context of the discussion in Utt and Fort (2002) related to the Gini coefficient of perfect competitive dominance in a sports league.
[62] See Utt and Fort (2002, p. 371) for further discussions about the Gini coefficient, the adjusted Gini coefficient and the NSQF ratio.
[63] The measure is originally based on a three outcome point score system.
[64] See chapter 3 in Groot for more about this measure. Originally, it is based on a three outcome system.

competitive balance for win dispersion. Player salary budgets and betting odds prior to the tournament are examples of the type of information sources that could be used to calculate ex ante competitive balance. Garcia and Rodríguez (2002, p. 20) use budget data in a demand function for football, related to both teams at the beginning of the season, "…. because they depend, among other things, on the salaries of the players, which should proxy their productivity".[65] Following up the discussion in Chapter 2.3.1, given unbiased betting odds, these should reflect market expectations of which team will win the championship providing a possible source of data to estimate ex ante competitive balance. Another possibility is to use measures based on calculated total quality index (TQI) (Gerrard, 2001, 2004a). The TQI is an index for team quality based on the player quality index (PQI), which is again based on a player's age, career total league games, career total goals scored, previous season appearances (number of games started, number of games as substitute, goals scored), and international recognition (full caps, U-21, last appearance). The estimation must be done before every season. This measure also takes into account players who changed teams the previous season (transfer date and the date for being able to play for the new club must be available). On basis of the TQI and eventually other relevant information, such as match schedule in a complex league context, one could predict the league outcome for the season ex ante level, i.e. ex ante competitive balance.

2.5.2. Performance Persistence

Apart from the dispersion of championship winners, early researchers in the economics of professional team sports were little concerned with performance persistence. This is reflected in the measures of competitive balance most commonly used. As Eckard (1998, 2001) points out, the variance in team performance from year to year in the league is not captured at all by measures such as the NSQF ratio.

Performance persistence concerns team performance over time, specifically the **correlation** from season to season, or as Scully (1995) calls momentum in sports. Daly and Moore (1981) and Balfour and Porter (1991) provide the first attempts to measure performance persistence. Daly and Moore (1981) apply the *Spearman's rank*

[65] In a related footnote, Garcia and Rodriguez (2002) refer to Falter and Pérignon (2000), including this kind of variable in a demand study.

correlation coefficient (SRCC) as a measure for testing changes in performance persistence in relation to introduction of free agency in the MLB. The following formula can be used:

$$SRCC = 1 - \frac{\sum_{i=1}^{N} d_i^2}{N(N^2-1)} = 1 - \frac{\sum_{i=1}^{N} (a_{i,t} - a_{i,t-1})^2}{N(N^2-1)}$$

where $(a_{i,t} - a_{i,t-1})$ represents the difference in ranking for team number i on basis of performance in season t-1 and season t, and N is the total number of teams. This measure is also applied by Maxcy (2002) and Maxcy and Mondello (2006).

Another measure for rank correlation coefficient is the Kendall's τ,[66] which is applied in the context of professional team sports by Groot (2008). According to Groot, Kendall's τ can for example be written like this:

$$\tau = \frac{2(P-Q)}{N(N-1)}$$

where P is the number of concordant pairs and Q is the number of disconcordant pairs. Given an order based on the rank between the teams in season t-1, the number of concordant pairs in season t is those in a natural order, and opposite for disconcordant pairs.[67]

Balfour and Porter (1991) calculate the *correlation* of teams' win percentage in season t with season t-1, t-2 and t-3. Significant correlations between team performance across seasons are interpreted as evidence of weaker competitive balance. Butler (1995) also uses correlation coefficients, but only considers first-order correlation between team's win percentages in two consecutive seasons. Gerrard (1998) uses correlation of league rankings and win percentages.

---

[66] I use τ instead of t, as in Groot, because it will not overlap with notations of other measures in this thesis.

[67] Following Noether (1981), a pair is concordant (disconcordant) when X_j-X_i and Y_j-Y_i have the same (opposite) sign, given that (X_i, Y_i) and (X_j, Y_j) are "a pair of (bivariate) observations" (p. 41). Further, see chapter 1.4 in Groot for calculating Kendall's τ between the first and the second season of the English top division (which had the same teams competing both seasons).

La Croix and Kawaura (1999) compare average win percentages over periods and relate persistence to changes between periods. For example, a low average win percentage in one period followed by a higher average in the next period might reflect an improvement in competitive balance.

Vrooman (1996) and Szymanski and Schmidt (2002) both use autoregressive regression models, measuring the relationship between team performance (win percentage) in one season (t) with that in the previous (t-1) season using data from several seasons. Ross and Lucke (1997) also apply an autoregressive model but of a higher order (t-2 and t-3).

A number of measures are related to *variation in performance across seasons*. Gerrard (1998) measures *teams' average position change* and *average absolute win ratio change*. In constructing two-dimensional measures, both Eckard (1998, 2001) and Humphreys (2002) use between-season competitive balance as a component of their measures for competitive balance. The measure proposed by Eckard (1998, 2001) uses each *team's variance over time against its average winning percentage*, while Humphreys (2002), as a part of his CBR (competitive balance ratio), in practice follows Eckard, but uses a *standard deviation approach* (Eckard, 2003). Scully (1995) calculates the *coefficient of variance* on basis of historical winning percentage for single teams over time, and next across these teams for the league.

Performance persistency among teams in sports might also be measured using methods applied to strategic groups and mobility barriers in studies of industrial organisation. Gerrard (1998) applies the *M-index* (originally proposed by Feigenbaum and Primeaux, 1983) to English football. Gerrard (1998) defines five different strategic groups for the English top division. The different groups are defined on the basis of standard deviations from the mean win ratio. The *D-index* is another measure of mobility (Gerrard, 1998). In general, the D-index measures the degree of change in market share over time. For English football, Gerrard measures the degree of change in win ratios from the previous season.

Looking at the level of performance from year to year, Gerrard (1998) anticipates that in a league with high level of competitive balance, "teams are less able to persistently achieve high levels of team performance" (p. 5). In other words, competitive balance can be reflected in the *performance level of the best teams from year to year*. If they typically reproduce high performance season after season, it might be an indication of weak competitive balance. Hence analysis of the performance of the best teams over time might say something about trends in competitive balance. Gerrard (1998) defines teams with win ratios greater than one standard deviation above the mean as high performing teams.

Ross and Lucke (1997) measure *the coefficient of variation* on basis of what they call *the incidence of competitive teams*, where competitive teams are five or less games behind the winner. During a period of T seasons, the number of seasons each team can be called competitive is counted. These numbers are averaged over all teams, and standard deviation is calculated. The coefficient of variation is computed by dividing this standard deviation on the average. If perfect (ex post) competitive balance, the coefficient of variation is zero. Further, Ross and Lucke (1997, p. 662) count the number of teams going from performing "bad to good" and "good to bad". Good performing teams are defined as either five or fewer games behind the winner, or as ten or fewer games behind the winner. They standardize this measure, comparing the actual number over a given period with the total number of teams (i.e. the percentage of "counts").

Other studies count the number of teams achieving a top k outcome during a T season interval. In a simple league context, the top k outcome is required to be less than N, where N is the total number of teams (that continuously participate) in the league. Szymanski (2001) uses k = 3, 5 and 10, and T = 3 and 7. Buzzacchi et al. (2003) count the number of teams that enter the top five over longer periods (10, 20, 30, 40 and 50 seasons) and develop a Gini type index related to the actual number of teams that capture a top k rank over a given period of time, relative to the idealized number. In the simple league context, the idealized number can be calculated by their closed league model.[68] One could also convert the measures Eckard (1998) and Sutter and Winkler (2003) use on college football (NCAA) in North-America to a simple league context by counting the number of teams to be ranked at top k over a given period of time, and next, calculate the concentration of top k teams by a HHI measure. Moreover, following Eckard and Sutter and Winkler, averaging seasonal *entry/re-entry* in the k rank could be calculated for an interval of T seasons, where entry/re-entry in the top k rank is defined for teams that have not reached this outcome over the last five seasons.

[68] For description of this model, see Buzzacchi et al. (2003) p. 176.

There are also a number of measures of performance over time that are only based on long-run averages. Hence, these measures say nothing about variations between seasons, and cannot therefore be seen as measures of performance persistence. However, if a team has a relatively low or high average win percentage, this might indicate high performance persistence. Examples of this kind of measure can be found in Quirk and Fort (1992) who calculate the *distribution of lifetime win percent*, summarizing each team's win percent over T seasons, and then divide by T. They also measure *standard deviation of the lifetime win percent*, and *compare the actual standard deviation with an idealized standard deviation*. Eckard (1998, 2001) measures the *variance of the teams' cumulative win percentage over a given period of time*.

2.5.3. Prize Concentration

In a simple league context, prize concentration is the concentration of championship winners. Several of the measures of win dispersion can also be used to measure the concentration of championship winners, such as *Lorenz curves* (see, for example, Quirk and Fort, 1992; Szymanski and Kuypers, 1999), *Gini coefficients* (Quirk and Fort, 1992; Fort and Quirk, 1995) and the *Herfindahl-Hirschman index* (Gerrard, 2004b). Further, Eckard (1998, 2001) uses the *relative Herfindahl-Hirschman index* to calculate the concentration of championship winners. Eckard (2001) measures the championship concentration, defining the team with the highest win percentage in the regular season as the champion. Eckard calculates the relative HHI as Actual HHI minus Expected HHI with Expected HHI calculated on the basis of every team having an equal probability of winning the championship every season.

Other measures of prize concentration include *counting the number of championships* for selected teams (Rottenberg, 1956) or all teams in the league (Quirk and Fort, 1992). Further, championship winners can be calculated as the *share* of either the number of years in the league (Quirk and Fort, 1992) or over a given number of seasons (e.g. Szymanski and Kuypers, 1999; Gerrard, 2004b). Buzzacchi et al. (2003) first count the number of different teams gaining the highest win percentage in any individual season over a given period of time, and then calculate a Gini type index. Dobson and Goddard (2001) use a measure of championship concentration based on the *distribution of* *championship points* with three points for winning the championship, two points for second place and one point for finishing in third position.

3. Competitive Balance in a Complex League Context

3.1. Introduction

Chapter 2 showed that even in the simple league context competitive balance is multidimensional. This chapter focuses on a real-world approach to league structures in the context of competitive balance. In general, these leagues are more complex than the league systems presented in Chapter 2. At least four main issues have to be taken into account when looking at differences between a simple league and more complex league systems. These are that rankings are not only based on matches won, that leagues might have a match schedule other than the (pure) round-robin system, that leagues consist of more than a regular season tournament, and the promotion and relegation system.

One of the characteristics of actual league systems is that they are generally multi-prize structures, which has significant implications for the notion of competitive balance. In this respect, the concept of competitive intensity will be introduced. This concept is one of the key outcomes of this thesis and is introduced to allow for the impact of the prize structure on the degree of competition in a sports league. The level of competitive intensity for a specified league is affected by institutional changes in league structures, such as changes in promotion and relegation mechanisms and qualification to UEFA's European club tournaments (i.e. the UEFA Champions League and the UEFA Cup).

The first part of this chapter will focus on the most typical elements that differ from a simple league system. The second part is based on prizes and competitiveness, including the concept of competitive intensity, and examines how the prize structure influences competitive balance.

3.2. Relegation-Promotion Systems

One of the assumptions in the simple league context is a closed league with the same teams competing against each other season after season. However, the team composition of real-world leagues tends to change over time due to reductions and expansions in the number of teams, as well as changes resulting from the relegation-promotion system.

The open league, organised as a merit hierarchy system with promotion and relegation, is an important aspect of the organization of many European team sports (see, for example, Szymanski and Ross, 2000).[69] A report from the European Commission (1998, p. 4) states that "the system of promotion and relegation is one of the key features of the European model of sport".[70] This way of organising team sports in Europe, where the best teams move up and the weakest teams move down the hierarchy on basis of sporting performance (even though exceptions happen through relegation as a consequence of administrative punishment), will lead to a natural hierarchy, where the best teams will operate in the top division, while weaker teams will "find" their place further down in the hierarchic system.[71] It also allows the possibility for newly established teams to enter at the lowest level and then be promoted to a higher level on the basis of sporting achievement. In principle, a team established today, starting at the lowest division in its domestic league, could, after a number of promotions, eventually qualify and win the UEFA Champions League. In this system it is difficult, even impossible, for teams to claim monopoly control over its territorial area. Noll (2002) emphasises that neither the Football League nor the Football Association in England recognize territorial rights. However, for both of these arguments, promotion and relegation systems are not necessarily the only requirements. For example, the professional league system in English football has a relegation and promotion system within the four tiers (open league), but earlier it was nearly closed when it came to automatic promotion into the English Football League from the next tier in the football pyramid. For example, Noll (2002) shows that the Football League in England was effectively a closed league between 1932 and 1950 with no team relegated from the bottom division (i.e. Division 4) during that period.[72]

The NAML are closed with changes in league composition only occurring by agreement of the current teams.[73] League memberships (and expansions) are "determined by the award of franchises" (Dobson and Goddard, 2004, p. 361), or as "the gift of the existing

[69] Andreff and Staudohar (2002, p. 39) use the expression "unique to European leagues". However, promotion and relegation is for example also a part of football in Latin-American leagues.
[70] See www.sport-in-europe.com/SIU/HTML/PDFFiles/EuropeanModelofSport.pdf. This is also emphasised in the 2004 Magazine for education and culture in Europe about the European Union and Sport (ec.europa.eu/dgs/education_culture/publ/pdf/mag/23/en.pdf).
[71] Fort (2000) claims that the system for franchise movements creates the same output in the NAML.
[72] See chapter two in Dobson and Goddard (2001) for more information about teams entering the English Football League.
[73] Fort (2000) argues that US College football has a relegation and promotion system, beside that qualification and not qualification for playoff could be regarded as a promotion and relegation form.

members" (Szymanski and Ross, 2000, p. 2). The NAML represent the highest level in a fixed hierarchy.[74]

Compared to a simple league, a promotion and relegation system will have three direct implications on competitive balance. First, on basis of sporting and financial incentive effects, is it hypothesised that, ceteris paribus, win dispersion will be better. Second, the turnover in league composition over time creates complications for comparisons of measures of cross-seasonal competitive balance. Third, the real-world complexity creates of the need to produce additional measures of competitive balance beyond those needed for the simple league context.

The key factor, when analysing differences between closed and open leagues, with regards to win dispersion, is the sporting incentives for weaker teams, and hence also financial effects. Relegation has, in general, both sporting and financial disadvantages compared to continuing in the (top) division. It is expected that the various stakeholders in a professional sports team, such as fans, owners, players, coaches etc., gain higher (sporting) utility from participating in the highest level division. Generally, the financial effects from relegation are expected to be negative due to reductions in the general level of gate attendances and the lower value of broadcasting and other image rights.[75] In a share price event study, Dobson and Goddard (2001) estimate the impact on English football teams of relegation and promotion. On average, they find that the next trading day's share price increases by about 22.5 percent if a team gains promotion to, or avoids relegation from, the FA Premier League. There is an average downward price adjustment of 15.6 percent when teams are relegated or fail to win promotion. The greater the financial difference between the tiers in a merit hierarchy, the stronger the expected incentive effects on team performance.[76]

Even weaker teams that are out of contention for the championship have incentives to improve their ability to win matches in an open league (see, for example, Noll, 2003; Szymanski, 2003b; Szymanski and Ross, 2000). Szymanski (2004) claims that the open

---

[74] This might be helped by franchise movement or threatening about moving. The financial strongest area will attract the teams in the NAML, and hence bring the major leagues into the highest level (see Fort, 2000).

[75] However, Noll (2002) shows that this is not necessarily supported empirically, applying data from the English FA Premier League (and First Division) from the 1990s.

[76] See, for example, Deloitte (& Touche) reports (such as the Deloitte & Touche annual review of football finance August 2000).

league system will "reduce the variance of within-season win percentages" (p. 32), and relates this to calculations of NSQF ratio by Kipker (2000), Forrest and Simmons (2002), and Szymanski (2003b). This ratio "seems much smaller in European soccer" (p. 32). Further, Noll (2002, p. 173) uses a "simplified model of a sports league, which is initially developed by El-Hodiri and Quirk (1974)", and shows that "the equilibrium spread in team quality in the top league is lower than it would be in a league of the same teams with fixed membership" (p. 175). This happens because teams in an open league, in which the higher division is more profitable, will, in general, seek higher quality than in fixed (closed) leagues. Therefore, one would, according to Noll (2002), expect win dispersion in the open top divisions in European football to be better than in the closed NAML, ceteris paribus.

In addition to the incentive effects, open leagues might, through their dynamics, also achieve better win dispersion compared to closed leagues (in a simple league context) as a consequence of the weakest team(s) continually being replaced by the potentially stronger team(s) from the second level. As long as the promoted teams are stronger than the relegated teams, the dynamic will improve competitive balance at the top division, ceteris paribus.

According to the general quality of leagues, this might have consequences for the expected win dispersion difference between open and closed (in a simple league context) leagues. According to Noll (2002), a system with extremely lucrative international matches, such as the UEFA Champions League, should adopt a relegation and promotion system. This fact is concluded by Noll (2002):

Hence, it makes sense to adopt a system that, all else equal, produces stronger teams. Thus, adoption of promotion/relegation is much like an arms race: Conceivable, all teams and top leagues would be better off without it, but once one nation adopts the system, the others have a financial incentive to follow.

Szymanski and Ross (2000) also emphasise the quality effect, and claim that a reason for this is that open league teams have greater incentives to invest in quality players than teams in a closed league. Further, they argue that better competition in the open league will increase consumer welfare compared to a closed league. However, it can be questioned if higher quality among the better teams only is driven by the relegation and promotion system above some threshold, if it is only the very high expected payoffs for better teams in post-seasonal tournament that provide sufficient incentives for investment in high quality squads.

Noll (2002) finds hypothetical departures from the expected win dispersion "advantages" for open leagues. One is the case of a promoted team that merely seeks to capture the temporary economic rent from membership of a higher quality league without investing in the higher playing quality required to have a realistic prospect of maintaining the team's status in the higher division. Such a team is likely to increase the degree of win dispersion in the higher division. Another is that if the top division consists of all the best teams, the promoted teams will, by definition, be weaker than the relegated teams, and hence will increase win dispersion. Note that these effects will only be temporary, since one would expect the relegated team(s) to be promoted the following season (or the promoted team improves its playing squad to be competitive).

The literature[77] comparing open leagues to closed leagues suggests that as well as having higher playing quality (see, for example, Noll, 2002; Szymanski and Ross, 2000), open leagues are likely to be weaker financially due to higher wage costs for the less competitive teams.[78] In addition, open leagues tend to have a smaller top division (Noll, 2002).[79] Some argue (see, for example, Andreff and Staudohar, 2002) that a cross-subsidisation policy in closed leagues is potentially more able to distribute revenues equally among the teams. However, the promotion and relegation system in European leagues does not necessarily preclude product and labour market restrictions although the incentives for the bigger teams in the bigger leagues to be competitive in the UEFA tournaments militate against such restrictions. Extensive cross-subsidisation is more likely in the smaller European football leagues. Sweden, for example, had a uniform distribution of revenues from broadcasting deals.

Promotion and relegation has an impact on uncertainty of outcome (see, for example, Sloane, 1971; Jennett, 1984; Szymanski and Kuypers, 1999). This is due to the

[77] According to Noll (2002) there has been little research on this field, mentioning only the work by Hall et al. (2002) and Szymanski and Ross (2000). Fort (2000) claims that many of the institutional differences between the structure in European and North American team sports are small in real.
[78] Noll (2003, p. 550): "...promotion and relegation system probably reduces profits, especially among the best teams in the top league. Promotion and relegation increase the value of the best players to teams that are in battle for either promotion or relegation."
[79] This is also claimed in Szymanski (2003b).

increased degree of match significance as well as the improvement in the win dispersion created by the incentive effects for weaker teams competing to avoid relegation. Szymanski (2004, p. 32) claims that a system with relegation and promotion also "...promotes uncertainty".

The above arguments on the competitive balance effects of promotion and relegation are mainly concerned with the win dispersion dimension. There has been little focus on the effects of performance persistence from the relegation and promotion system, other than as suggested explanations of empirical results (Szymanski and Smith, 2002; Buzzacchi et al., 2003). Intuitively, weaker teams might reduce their probability of relegation by over investing in players and operating with a wage-revenue share above a viable limit. In other words, such teams "play lotto" where the prize is to stay in the division and the downside is to be relegated with financial problems. This type of behaviour is likely to increase the volatility of team performance over time, and hence reduce performance persistence.

Implications for Existing Competitive Balance Measures

The NSQF ratio can be applied in an open league system, without modification, because the measure is only related to within season outcomes, and is hence not tracking the performance of specific teams across seasons. Since this measure allows for changes in the number of matches played in the league, it is robust to expansions and reductions in the number of teams in the league over time. It can therefore both be used to compare competitive balance across leagues and within leagues over time. On the other hand, there are major difficulties in measuring performance persistence and prize concentration in a multi-divisional open league as teams move between divisions and may compete for different prizes in different seasons. Measuring prize concentration in lower divisions is particularly difficult since, by definition, winning the prize (i.e. promotion) means that the team will not be competing for that particular prize in the following season. This is one reason why studies of competitive balance in European football have tended to focus only on the top tier. The general problem of converting closed league competitive balance measures to open leagues is the higher frequency of alternation of teams in the top division in an open league. In addition, relegated teams move into a "new" division. Directly transferring competitive balance measures applied in closed leagues, especially performance persistence measures, to open leagues, can hence be inappropriate. For example, the components of the measures by Eckard (1998, 2001) and Humphreys (2002) relating to performance variation for a given team over time are very difficult to transform for use in open league systems.

Both *Spearman's rank correlation coefficient (SRCC)* and the *Kendall's* τ are based on the continuous participation of the same teams in a repeated tournament. They therefore have to be modified in a system with relegation and promotion (see also Groot, 2008). For a given season, they measure the correlation between the current and previous season rankings. In a closed league (without expansions/reductions), the teams competing in the previous season (t-1) are exactly the same as in the current season (t). On the other hand, in an open league system, a number of the teams at the bottom of the standings at the end of previous season are replaced by promoted teams prior to start of the current season. Because of the changes in divisional team composition across seasons, these measures must be modified for application in open leagues. One approach is to replace the teams relegated at the end of the previous season by the promoted teams in the rankings of top division for the previous season. This approach is illustrated in the following example, using a ten team league with the bottom two teams automatically relegated at the end of each season:

Season t-1	**Modified season t-1**
Top Division	
1. Team 1	1. Team 1
2. Team 2	2. Team 2
...	...
...	...
8. Team 8	8. Team 8
9. Team 9 – **Relegated**	9. Team A - **Promoted**
10. Team 10 – **Relegated**	10. Team B - **Promoted**
Second-Level Division	
1. Team A – **Promoted**	
2. Team B – **Promoted**	

For each season, for example the SRCC is calculated as the correlation between the end-of-season rankings in the current season, with the modified rankings for the previous season with the newly promoted teams treated as if they had finished bottom of the top division in the previous season.[80] However, promotion and relegation procedures differ

[80] I prefer this ex ante ranking of the promoted teams instead of using the ex post procedure in Groot (2008). Groot's internal rankings between the promoted teams (from season t-1) are a function of the final

between leagues (as well as within leagues over time). The worked example above is only applicable to leagues with automatic relegation. Some leagues, such as the Norwegian, use a combination system with the bottom two teams relegated automatically but the third bottom team qualifies for a relegation/promotion playoff together with the third best team from the second level division. If the league above has a similar system, the example above can be modified as follows:

<u>Top Division</u>	**<u>Second-Level Division</u>**
8. Team 8 - **Play-Off**	1. Team A - **Promoted**
9. Team 9 - **Relegated**	2. Team B - **Promoted**
10. Team 10 - **Relegated**	3. Team C - **Play-Off**

If the third team in the second-level division wins the relegation promotion playoff, it is suggested that this team is treated as the lowest ranked team in the modified season t-1 rankings:

8. Team A - **Promoted**
9. Team B - **Promoted**
10. Team C - **Promoted via Play-Off**

An alternative method to deal with the relegation and promotion system in the calculation of for example the SRCC is to eliminate the promoted team(s) in the(ir) first season in the top division.

Expansions and reductions of teams in an open league can also be dealt with by both of the alternative methods above. It might be more difficult to deal with expansions in closed leagues, such as the NAML. Three alternatives might be considered:

1. The method used by Daly and Moore (1981), which is to exclude the expansion teams from the calculations. They only analyse performance persistence by calculations based on the "original teams".
2. As in the alternative method for dealing with a promotion-relegation league above, exclude the performance of these teams in its first season (when they have no previous season ranking) but include them thereafter.
3. Use a predicted previous season ranking for the expansion teams. One obvious assumption is to treat the expansion teams as the weakest teams. This is consistent

standing in season t. In the example above, team A could be ranked either nine or ten in Groot's "modified season t-1", based on the final table in season t, in comparison with team B.

with historical experience of expansion teams who generally perform poorly in their first seasons (Eckard, 2001). If the expansion occurs because of the merger of two previously independent leagues, there are two possibilities. For example the SRCC can be calculated separately for the teams in the two constituent leagues after the merger, by continuing to rank the teams as if they formed two separate leagues. Alternatively a single composite ranking of the teams for the seasons prior to the merger can be created using win-ratios to allow the calculation of a comparable combined-league SRCC before and after the merger.

As shown in the previous chapter, Gerrard (1998) uses a number of correlation measures to capture performance persistence. These measures are applied in an open league (English football) by ignoring newly promoted teams. This is done for the measures of average positional changes, average win ratio changes and the M-index, while both relegated and promoted teams are included in the calculations of the D-index.

The Gini coefficient measure used by Buzzacchi et al. (2003) is also constructed for an open league system. The idealized number of top k ranks must be modified to allow for increases in the number of teams in the division over time.

Specific Measures Constructed for Open League System

Another approach to the analysis of competitive balance in a promotion and relegation system is to construct new measures specifically focusing on the teams involved in promotion and relegation. If there is perfect competitive balance in the top division, the newly promoted teams should have the same probability as the other teams in the division of achieving any given sporting outcome. The alternative hypothesis is that newly promoted teams are competitively weaker. Analysis of the performance of promoted teams can shed light on trends in the difference in sporting performance between tiers. This is very topical because European football leagues are increasingly concerned with the widening financial gap between teams in the top and lower divisions and the existence of so-called "yo-yo" teams.[81] A greater frequency of these teams might indicate an increased gap between the two tiers, so that it is more difficult for newly promoted teams to survive in the higher division. This kind of analysis extends

[81] Teams being too strong to stay in the lower division, but too weak to survive in the higher division.

the "normal" use of the competitive balance concept in European football, where it is the top division in the league that is usually analysed.

Gerrard (1998) suggests a number of measures relevant for promotion and relegation, such as comparisons of *the length of tenure of teams in the top division, the average win ratio in the first season in the top division for newly promoted teams, the relative performance of promoted and established teams,* and *the successful newly promoted teams over a given period of seasons*, where successful is defined as win ratio > 0.5. In a perfectly balanced league, newly promoted teams should have the same probability of a win ratio above the mean as teams continuing in the top division.

Frick and Prinz (2004) measure the *relative performance of promoted and established teams* by comparing average points scored for newly promoted teams with the average points for the other teams in the top division, and similarly, by putting a variable for newly promoted teams as one of the explanatory variables in a fixed effect panel data model, where point score (using league rankings yields similar results) is the dependent variable.

The growing frequency of yo-yo teams may be measured by *the survival rate for promoted teams* (Gerrard, 1998) compared to the relative frequency of newly promoted teams remaining in top division after *n* seasons (see also Frick and Prinz, 2004). Further, Frick and Prinz construct a fixed effect logit model with relegation as the limited dependent variable and being a newly promoted team is one of the independent variables. They also use a Cox proportional hazard model, analysing survival of promoting teams through having "the spell duration" (i.e. the number of season in the division since the team's most recent promotion) as the dependent variable with a number of independent variables including being a newly promoted team.

An alternative measure is an *n season relegation ratio for promoted teams*. This generalizes the measure of survival rate for promoted teams by measuring the actual survival rate relative to a theoretical benchmark of the idealized rate of relegation of newly promoted teams under conditions of perfect competitive balance when both newly promoted and continuing teams have the same probability of being relegated. For example, for the first season after promotion, the relegation probability for all teams should be R/N, where R is the number of teams to be relegated, while N is the total number of teams in the division. Expanding to a two-season period, the relegation probability for the second season, with regards to the probability for teams to stay in the division during the first season, is [1-(R/N)]*(R/N). Below, the ratio measure is the actual number of promoted teams that within two seasons are relegated, against the theoretical benchmark, and can be defined as:

Two season relegation ratio for promoted teams
= Actual number / $\{(R/N)[2R-(R^2/N)]\}$

This formula is a simplification, where the number of teams to promote is equal to the number to be relegated (R is both related to the number of teams to promote and to relegate). However, empirically, the formula can easily be made more complex to capture both differences in the number of teams to promote and to relegate, as well as playoff systems related to promotion and relegation.

If there is perfect competitive balance, one could also expect that over a longer period of time, newly promoted teams should have the same relative frequency of post-seasonal qualification and championship titles as continuing teams. For example, the English FA Premier League of today consists of 20 teams. This means that if this league was perfectly (ex ante) competitive balanced, each team has 1/20 probability of winning the championship. Since three teams are promoted each season, the expected probability for a newly promoted team to be the champion should be 3/20. The same principle can be used for qualification to the UEFA Champions League. Currently, the top two at the end of the season are automatically qualified for the UEFA Champions League, while the teams finishing third and fourth enter the UEFA Champions League 3rd qualifying round. In a perfectly balanced league, each team should have 1/5 probability of reaching the top four, and hence there is a probability of about 51 % that one of the top four should go to a promoting team, implying that on average, every second season a newly promoted team should finish in the top four in the league.[82]

3.3. Rankings Not Based Solely on Matches Won

Many leagues have systems for valuing results that are not only based on matches won. One major difference between leagues is whether or not the score systems allow for tied

[82] Calculated on basis of [1 – (17/20 * 16/19 * 15/18 * 14/17)].

(or drawn) matches. Tied matches are typically allowed in the various codes of football (except American Football) as well as ice (and field) hockey and handball.

Leagues allowing tie matches need to extend the valuation system beyond that allowed for in the simple league context. Historically, European (association) football has mainly used two different systems for valuing draws. The (2,1,0) score system was adopted by the Football League in England at its start in 1888 with wins awarded two points, a draw treated as a "half-win" with one point to each team, and a loss valued as zero. The current (3,1,0) score system with three points for a win gives a relatively lower value for draws (1/3 of a win compared to 1/2 of a win). However, there are several cases in European football of deviations from these two main systems.[83]

There has been little research of the effects on competitive balance from changes in the point score system in a league, including the effect of allowing for tied outcomes. However, there are several studies of the impact of changes in incentive effects, for example overtime play in the NHL (Abrevaya, 2004; Easton and Rockerbie, 2005; Shmanske and Lowenthal, 2007). Easton and Rockerbie (2005) and Shmanske and Lowenthal (2007) also analyse the effects of whether matches are between teams in the same or different divisions and conferences. Shmanske and Lowenthal find that when playing against teams from same conference, there is an increased probability of a tied game (statistically significant for the home team, but not significant for the away team). Recently, Haugen (2008) has analysed theoretical effects of changes from the (2,1,0) system to the (3,1,0) system on basis of game theory both related to playing strategy and competitive balance.[84] His theoretical results suggest weaker competitive balance after changing point score system.

The other main issue in the context of competitive balance is how to treat a drawn outcome in comparisons between leagues with and without draws. In general, win dispersion is expected to be lower in leagues with draw outcomes, compared to leagues with only win-loss outcomes, ceteris paribus. If a match has to be played until a winner is decided, with the same form of play as in regular time (i.e. no shoot-outs or other constructed tie-breaker mechanisms), the better team will have a higher probability of winning the match than the weaker team. Hence, one would anticipate the "final result"

[83] See, for example, www.rsssf.com.
[84] Based on game theory. Brocas and Carrillo (2004), among others, also discuss playing strategy in relation to values of wins in football (three versus two points).

distribution of "previous tie matches" to be skewed (towards the better teams) if there is overtime play. This follows Cain and Haddock (2006) who analyse major league baseball where the game continues indefinitely with extra innings until one team wins. It would be expected that the NSQF ratio should be lower in leagues with tie matches.

There is a number of possible ways to treat draws in win dispersion measures. One approach is to use point percentage. In the case of the (2,1,0) scoring system, win and point percentages will be equal (Cain and Haddock, 2006; Fort, 2007). Under other scoring systems, such as the (3,1,0) system, point percentage will capture the relative value of wins and ties correctly. An alternative approach is to ignore tie matches, and hence calculate win percentages only on the basis of matches won or lost. The idealized standard deviation will then need to be calculated using the number of "relevant" (i.e. won/lost) matches which will be lower than the actual matches played. A third approach, which seems to be the most popular in analysing European football, is to ignore the different values of wins, ties and losses, and transform them into a (100%,50%,0%) system (i.e. win percentage, where ties are treated as half wins). In this case, the actual standard deviation is not affected by changes in scoring systems. More general, this case can be extended to a given share of wins. A fourth possibility is to relate measures of the actual point structure (Cain and Haddock, 2006; Fort, 2007).

Another possible complication is that the valuation of drawn matches might be dependent on a period of overtime when teams are tied at the end of regular time. Leagues may weight wins achieved in overtime differently from wins in regular time, such as in the top division in the Norwegian ice hockey league (www.hockey.no). Leagues also differ in how they treat byes when there are an odd number of teams, implying that one team has no scheduled match in each round of matches. Normally leagues ignore byes in the calculation of points, ensuring that every team has the same number of byes. However, the National Rugby League in Australia awards two points for a bye, the same as for a win. Although this does not affect the league standings (since all teams have the same number of byes in a season), it affects the calculation of win dispersion measures and creates difficulties when comparing competitive balance across leagues.

In summary, compared to the win-loss outcome in the simple league context, real-world leagues have two main issues to further take into account when it comes to determining rankings: the range of match of outcomes and the relative values of these outcomes. These issues have been discussed recently by Cain and Haddock (2006) and Fort (2007) in relation to the NSQF ratio.

Multiple Outcomes and the NSQF Ratio

The NSQF ratio was originally designed for leagues without (or with very few) tied matches. It is based on win percentage, where a win is equal to 1 (100 %) and a loss is equal to 0 (0 %), and the expected outcome of each match is 0.5 (50 %). This is equal to the definition in the simple league context. In other words, the simple league NSQF ratio is related to the following valuation system:

$$NSQF = \frac{ASD(100\%,0\%)}{ISD(100\%,0\%)}$$

where ASD is the actual standard deviation in a league that is based on two outcomes; win or loss, and where a win is valued 100 %. The comparable system is given in the denominator. Usually, one would have named it ASD(1,0), but because of the recent discussion in Cain and Haddock (2006) and Fort (2007), it is appropriate for this thesis to use the values in percent, to divide it from points.

When including the NHL in their analyses, Quirk and Fort (1992) have "transformed" this measure to also include leagues where ties are more usual (even if it is less usual than in European football in general). Measures based on win percentage/win ratio have typically treated tied matches in European football as half-wins (Cain and Haddock, 2006). This seems straightforward for the (2,1,0) system and can easily be converted to a (100%,50%,0%) system. On the other hand, Cain and Haddock (2006, p. 331) claim that this "is generally inappropriate", and show that the NSQF ratio is higher when applying the three point score into the calculation of the ratio (because lower ISD).

Calculating the actual standard deviation in a (100%,50%,0%) system is straightforward. The potential problem for the NSQF ratio in leagues with tied games is how to calculate the idealized standard deviation. Typically, the ties are ignored in the idealized standard deviation in analyses of the competitive balance ratio in European football. This creates a potential problem, since the scope of the measure in the actual standard deviation differs from the idealized, as can be shown in the following formulas:

$$NSQF_{No-draws} = \frac{ASD(100\%,0\%)}{ISD(100\%,0\%)}$$

$$NSQF_{Draws} = \frac{ASD(100\%,50\%,0\%)}{ISD(100\%,0\%)}$$

However, although this is not a problem when comparing the NSQF ratio between European football leagues, it does create difficulties for cross-Atlantic comparisons between European football leagues and NAML (at least the MLB, NBA and NFL) since the way tied matches are treated in the idealized standard deviation will affect NSQF ratio comparisons. As mentioned earlier, the potential problem is that drawn outcome possibilities should reduce the expected distribution of outcomes, and hence give these leagues a lower NSQF ratio than if draws were taken into account in the idealized standard deviation:

$$NSQF = \frac{ASD(100\%,50\%,0\%)}{ISD(100\%,0\%)} < NSQF = \frac{ASD(100\%,50\%,0\%)}{ISD(100\%,50\%,0\%)}$$

The NSQF ratio might be seen as overestimating the win dispersion in European football, because it ignores one of the outcomes and thus underestimates the idealized standard deviation. If an idealized league (i.e. perfect competitive balance) is conceptualised as a uniform distribution of match results (i.e. equal probability of win, tie and loss) then the alternative idealized standard deviation will be $\frac{1}{\sqrt{6m}}$.[85] This tie adjusted idealized standard deviation is smaller than the original idealized standard deviation with two outcomes in the NSQF ratio.[86]

A uniform distribution of the three outcomes is arbitrary, but can be justified as representing maximum uncertainty of match outcome and hence be consistent with perfect competitive balance. However, this might not be a straightforward case, because

[85] Expected outcome is 0.5. Var(1,0.5,0) = 1/3x(1-0.5)^2 + 1/3x(0.5-0.5)^2 + 1/3x(0-0.5)^2 = 1/6.
[86] See Fort (2007) for more about the relationship between the idealized standard deviation, the NSQF ratio and different point score systems. This will also be discussed later in this section.

the expected outcome of a match with three outcomes in a league, where all teams are of equal playing strengths, could very well be argued to be tied (Bourg, 2004). In this discussion one might also take into account that there are significant differences in scoring systems/scoring frequencies between different sports. The latter will lead to a greater number of non-tied results in those sports, where the number of scores is much higher than in, for example, European football, even if the teams in both sports could be of approximately equal ex ante sporting quality. In general, there are two issues in an idealized (perfect balanced) league in this setting:

1. The expected probability of wins should be equal to the probability of losses, and hence that in a (1,0.5,0) score system, the expected outcome of each match is 0.5 (a draw when three outcomes).
2. The expected probability of a tie is difficult to decide. It is dependent on the distribution of different scores. Given a normal distribution of the scores, with $E(x) = 0.5$, it must be the size of the 0.5 area that determines the ex ante probability of tied matches. On one side one can argue, as above, that all three outcomes should be equal. However, on the other side, if there is equal sporting quality, although a tied match is expected, matches could also end as wins/losses within a given confidence interval. The importance of stochastic variation in match outcomes is in part a function of the game structure in different sports.

An alternative approach is to use the simple league context as base, where each team has 50 % probability of winning and 50 % probability of losing. When team A and team B play against each other, the following outcomes from the different states are given:

$$\{W_A W_B, W_A L_B, L_A W_B, L_A L_B\}$$

Since an equal outcome for both teams results in a draw, the following results are given: $W_A W_B$ = draw, $W_A L_B$ = win for A, $L_A W_B$ = loss for A, and $L_A L_B$ = draw. This means that the probability distribution for team A, when playing team B, is 25 % on win, 50 % on draw, and 25 % on loss.

Cain and Haddock (2006) have a different approach to the theoretical problems of including tied matches into the measure of idealized standard deviation, claiming that (p 331): "...the theory of equal ignorance does not generate a probability before the fact for a win-loss versus a tie. That is an empirical question". In other words, Cain and Haddock (2006) use ex post calculations to determine point 2 above. They calculate that about 25 % of the matches in the two highest level divisions in English football are ties, and therefore come up with an alternative idealized standard deviation in the NSQF ratio,[87] on the basis of (3/8-2/8-3/8) probabilities to $\sqrt{.75 \times m}$, given the (2,1,0) point system.[88]

Cain and Haddock (2006) solve the "tie problem" by calculating the percent of matches that actually have ended as ties. However, to justify the use of the idealized standard deviation, the (ex ante) perfect competitive balance should be used as base. The actual number of tied matches cannot be more than an indication of the ex ante outcome probabilities. One problem is that the share of tied matches varies from league to league and across seasons. A relevant question is therefore, why is the English top and second level division used to find the "proper" share of tied matches? The English top division is not the most competitively balanced in Europe (see Chapter 4). Would they have used a different percentage if they had used the Italian league, where the draw share is more than 30 percent, as base, or do they mean that different percentages for different leagues should have been applied? Another feature of the English top division is that the percentage of tied matches is higher after the Second World War (25.9 %, own calculation) compared to that over the whole history of the division (24.59 %, source: Cain and Haddock, 2006). Would they have used 26 % if their analysis was only based on post Second World War data? This is arbitrary. If it is necessary to draw-adjust the idealized standard deviation, one should apply a less arbitrary approach, so it could be possible to apply the measure in similar leagues. Therefore, the statement about theoretical and empirical issues, when it comes to tied matches in Cain and Haddock, has to be questioned. Is it not problematic to apply observed results in a measure for perfect competitive balance, when we know that leagues typically are not perfectly balanced? On the other hand, using observations, such as in Cain and Haddock (2006), can be relevant for indicating the relevance of tied matches. Based on my calculations, less than ten percent of matches end as draws in handball in Norway and Denmark.

---

[87] As is confirmed in Fort (2007), Cain and Haddock (2006) use the idealized standard deviation both about what Quirk and Fort (1992) denote the idealized standard deviation and what this thesis calls the NSQF ratio. In table 7.1 in the 1992 edition, Quirk and Fort explain that the Ratio = Actual Standard Deviation/Idealized Standard Deviation. See Fort (2007) for further discussions.

[88] Note, as is described in Cain and Haddock, that the idealized standard deviation is differently calculated under a point score system (multiplying with $\sqrt{m}$) than if percentages are used (dividing on $\sqrt{m}$). This is followed up later in this section. Var(2,1,0) = $3/8x(2-1)^2 + 2/8x(1-1)^2 + 3/8x(0-1)^2 = 3/4$.

Hence, it is more likely that a European football match ends as a tie than a handball game.[89] Taking these relations into account, deciding the idealized standard deviation might be even more difficult, and perhaps the method by Cain and Haddock is the better solution. However, the expected probability for ties in a team sport might be generated from a function that consists of variables, such as the length of the match, the score frequency, the characteristics of play, and so on, into a set of score probabilities.[90]

Since the methods above satisfy the requirement that the probability of wins and losses are equal, none of them can be rejected on logical grounds as possible measures of idealized standard deviation in a three outcome league. "Absolute" perfect win dispersion is different from the "expected" perfect win dispersion within a normal distribution. One might argue that in a league with "absolute" perfect competitive balance, all games should end as a draw. However, even if all teams are equal in sporting quality, luck, critical incidents, the structure of the matches etc., will affect results, and these factors should be more relevant the more limited time there is to make the final score (such as Cain and Haddock discuss in the context of baseball). Therefore, one would expect that even in a situation of perfect "sporting quality" across the teams, many matches will not end as ties. Using the normal distribution with the draw as the mean outcome seems most appropriate, but when is the deviation from the expected value big enough to result in a win/loss outcome? In other words, what is the expected number of drawn matches in a league that consists of teams of equal sporting quality? Given a realistic range of possible match outputs, the predicted number of wins/losses in sports may be highest, even if the individual probabilities for drawn matches are high.

A difference between the original idealized standard deviation in the NSQF ratio and an alternative idealized standard deviation that aims to capture the possibility of tied matches is that the win-loss idealized standard deviation only has one point reflecting maximum win dispersion, which is when the number of wins is equal to the number of losses for all teams in the league (assuming that all teams play the same number of games). In a situation that allows for tied matches, there are multiple possible levels of maximum points depending on the number of tied matches, since all situations where the number of wins is equal to the number of losses will give maximum win dispersion.

[89] According to Koning (2000), Stefani (1983) analyses predictabilities of match outcome across different sports.
[90] Further discussion about probabilities for draw matches is beyond the scope of this thesis.

Cain and Haddock argue that probabilities cannot be determined theoretically. But as argued above it is doubtful if it is appropriate to use empirical calculations in a "tie adjusted idealized standard deviation". The question remains: is it necessary to find a "tie adjusted idealized standard deviation" for comparing competitive balance ratios between different types of leagues? Is it enough to be aware of the effects on competitive balance of three outcome matches, compared to the traditional use of the NSQF ratio? A starting point is to consider the two uses for the NSQF ratio:

1. A measure for comparing win dispersion within a league over time and/or between leagues, where different numbers of matches is taken into account.
2. A measure for indicating the deviation of the actual sporting outcome from the expected outcome of a league, consisting of teams having equal playing strengths (i.e. ex ante perfect win dispersion given balanced match schedule[91]). The NSQF is therefore a measure to indicate the level of the sporting quality balance.

These two points seem to be used indirectly in Cain and Haddock's descriptions of the tie-adjusted NSQF ratio. Related to the first point they claim that (p. 331): "...it becomes especially problematic for cross-league comparisons where one league (such as soccer) experiences a large number of ties while another (such as baseball) plays each game until there is a winner". They also relate their discussion to the second point where the NSQF ratio (p. 331): "...seems unobjectionable as a measure of parity. A subtle difficulty arises, however, if their measure is used to examine parity within a league that awards points to ties...".

The NSQF ratio is excellent for the first purpose, regardless of league structure. The ratio is constructed as a measure that can be used to compare competitive balance in different leagues and within leagues with different number of matches played during time. In other words, there is a need for a "benchmark" to measure against when doing competitive balance comparisons, in line with, for example, financial ratios when

[91] If perfect (ex ante) sporting quality balance, it is expected ex ante perfect win dispersion regardless of playing schedule. Therefore it might look unnecessary to require balanced schedule. The problem is that perfect ex ante win dispersion might also be achieved for a league without perfect ex ante sporting quality, but as a consequence of unbalanced match schedule. For example in a league with eight teams, where four of them are much better than the four other teams, and the (ex ante) sporting quality within the two groups are equal, perfect (ex ante) win dispersion should be expected, if the matches are only played between teams within the two groups.

comparing financial performances between companies. Using a simple league as the base for this benchmark is one solution. No league satisfies all the attributes of a simple league system, which is not the aim. The purpose is to measure how far away the different leagues are from "the benchmark", and then apply these results for comparisons across and within leagues over time.

The second purpose of the competitive balance ratio is to calculate how far away these leagues' actual competitive balance is from what should be seen as the idealized competitive balance given perfect (ex ante) quality balance. It is doubtful how well this purpose can be satisfied by the NSQF ratio without several adjustments, even in leagues with binominal outcomes. There are many ways in which leagues are organized, and leagues in general deviate from at least one of the attributes of a simple league system. These deviations might therefore affect the relationship between the distribution of sporting quality between the teams in a league (ex ante sporting quality balance) and the distribution of sporting outcome. Applying the (ex post) distribution of sporting outcome to indicate the (ex ante) sporting quality balance in a league is much more complex than that reflected in the idealized standard deviation part of the NSQF ratio. Tied matches, score systems, playing schedule, incentives (transfers and risk-taking behaviour as regards expenditure on playing talent) created by the possibility of being relegated, incentives in relation to drafting, and incentives after secured playoff place are examples of relevant factors that should have been taken into account. It therefore looks very difficult to apply the NSQF ratio as a perfect measure for comparing the sporting quality balance across leagues (but it will of course give a useful indication). This is confirmed in Utt and Fort (2002) when constructing a theoretical perfect competitive dominance measure that takes into account factors such as tournament structure.

However, the latter method might provide an alternative for reflecting the sporting quality balance, using a benchmark for imperfect competitive balance, as is done in the Goossens' index (Goossens, 2006) and in the adjusted Gini coefficient by Utt and Fort (2002). These measures will be affected by drawn matches in the "actual" part of the calculation, but not in the "idealized" part. However, Utt and Fort (2002) show there are still many similar difficulties.

Differences in the level of restrictions on product and labour markets have received most attention as an explanation of *why* win dispersion differs between leagues. This thesis seeks to expand the list of possible of determinants to include league structure, and the possibility for tie matches. Tied matches will reduce the expected standard deviation compared to the idealized (benchmark) league. Similar effects might be found in the structure of the NAML, where leagues are divided into different conferences and the match schedule deviates from a full round-robin schedule. For example, in an unbalanced-schedule league, where the stronger teams are scheduled to play more frequently against other stronger opponents, a better competitive balance would be expected. The possibility of tied matches is not the only structural issue that affects the expected (end-of-season) competitive balance in sporting leagues.

In the NSQF ratio it is the distribution of sporting outcome (win dispersion) in the league that is the focus. When this distribution is better in European football than in the NAML, adjusted for the number of games played, the NSQF ratio should capture this, because it is actually this distribution that is of interest to most stakeholders. If the standard deviation of sporting outcome is better in Europe, this should be reflected in the ratio, even if tied matches are one of the determinants. By adjusting the idealized standard deviation for the possibility of tied matches, the better distribution of actual sporting outcome will be absorbed into the idealized standard deviation. This may be misleading because it adjusts for only one specific structural deviation from the simple league context, while other leagues may have other structural differences that affect competitive balance, but are not being controlled for in the modified measure.

In summary, the original NSQF ratio seems to be an appropriate solution as a measure for comparing end-of-season win dispersion across leagues, and over time within leagues. The effects of tied matches on the sporting outcome is reflected rather than ignored in the NSQF ratio. However, the ratio might not be perfect as an accurate measure of sporting quality balance (but is appropriate as an indicator).

Cain and Haddock (2006) also focus on the scale of the valuation system. This is followed up in the discussion by Fort (2007). Motivated by the point score systems in, for example, European (English) football, a number of scenarios are applied to show that the NSQF ratio is affected by these issues. They both demonstrate that, for example, a NSQF ratio in the (2,1,0) point system increases the ratio's value, compared to converting it to a (100%,50%,0%) system. On the other hand, Fort (2007) shows that NSQF is equal when using percentage of both (1,0.5,0) and (2,1,0), and he also claims that the results from the percentage converted (3,1,0) are close to similar. In addition, the correlation between the measures, both based on percentages and absolute points, is close to one. Fort finds it difficult to choose a preferred system when it comes to calculations of the NSQF ratio, at least for analyses over time, because of the nearly perfect correlation between the different approaches. However, when it comes to fan expectations and NSQF ratio as a measure for uncertainty of outcome, Fort emphasises that (p. 648-9):

> ....the "right" measure is the one that most precisely captures the impact of outcome uncertainty on fan demand; both the level of that uncertainty and its behavior over time are important to that end (Fort & Maxcy, 2003). For fan demand estimation, then, there is a justifiable suspicion that the (1,0.5,0) approximation only imprecisely captures the impact of outcome uncertainty on fan enjoyment and, hence willingness to pay.......A similar suspicion seems justified when addressing the incentives inherent in, say, moving to (3,1,0) as a replacement for (2,1,0)........But once this literature moves on to the impacts of different point allocation schemes on fan demand, absolute point versions of RSD may more precisely capture fan enjoyment.

According to Fort, if absolute point score was included in the calculations for the NHL, the NSQF ratio would have been 17 % higher, and for European football 15 % higher. Fort (p. 650) emphasises that this is the contribution of Cain and Haddock:

> But that is about all that can be said given the C-H article. Over time, absolute point measures and percentage measures of RSD all yield identical changes. So there is nothing in the C-H article to suggest that previous finding about the behaviour of competitive balance over time using any measure of RSD need to be revisited.

It is worth looking more closely at the assumptions behind some of the calculations in the articles by Cain and Haddock (2006) and Fort (2007). I disagree with Fort's (2007) equation (1), when referring to Cain and Haddock (2006) page 332, ISD(1,0.5, 0) = $\sqrt{0.25 \times m}$, because it seems that he is applying the binomial case mentioned in Cain and Haddock. Following the ex post match outcome percentages applied in Cain and Haddock, I suggest that the ISD(1,0.5,0) in Fort should have been:

$$\text{ISD}(1,0.5,0)_{\text{adj}} = \sqrt{(3/16)\times m}$$

because $(3/8)\text{x}(1-0.5)^2 + (2/8)\text{x}(0.5-0.5)^2 + (3/8)\text{x}(0-0.5)^2 = 3/16$.

This is important, because it can now be shown that there are no scale problems related to the NSQF ratio, as long as scales are similar among the different outcomes.

Proof: Since ISD(2,1,0) = $\sqrt{(3/4)\times m}$, it follows that the scale differences between ISD(2,1,0) and ISD(1, 0.5, $0)_{\text{adj}}$ is:

$$\frac{\sqrt{(3/4)}}{\sqrt{(3/16)}} = 2$$

Fort shows that ASD(2,1,0) = 2.000 x ASD(1,0.5,0). Hence is NSQF(2,1,0) = NSQF(1,0.5,0).

This is also robust for other combinations of win, draw and loss percentages as well. For example 1/3 probability for each outcome will give ISD(2,1,0) = $\sqrt{(2/3)\times m}$, while ISD(1,0.5,0) = $\sqrt{(1/6)\times m}$, which means that ISD(2,1,0) = 2 x ISD(1,0.5,0), and hence that NSQF(2,1,0) = NSQF(1,0.5,0).

In his equation 9, Fort (2007, p. 646) suggests a constant relationship between NSQF(3,1,0) and NSQF(1,0.5,0). But the given relationship is difficult to predict because there are different scales of the values involved. While the wins are tripled in absolute value, the absolute value is only doubled when it comes to draws. The relationship between the two ratios will therefore vary by the share of matches ending as draws and wins for the different seasons. This problem can be illustrated by calculating the scale differences in Table 2 of Fort (p. 647). For the English Premier League in 1997/98, the (3,1,0) score system is about 16 % higher than for the (1,0.5,0) score system, while the difference is only 10.4 % in the next season.

It is worth taking into account that the calculations in the examples by both Cain and Haddock (2006) and Fort (2007), when comparing different solutions for the first season of the English football league, show differences in point score systems and point percentage systems. These articles focus on differences in the NSQF ratio, without explaining why they appear in the calculation (Table 1 for Cain and Haddock and Table 1 in Fort). There is a logical reason for the differences, because both articles *include* draws in the idealized standard deviation for points, but *not* when calculating the NSQF ratio for the point percentage. If the latter group of calculations had also included draws in the idealized standard deviation, the NSQF ratios would have been equal. In addition, Fort is also differentiating in the use of draw in the idealized standard deviation, when comparing (1,0.5,0) and (2,1,0) point score systems, where the former is calculated without draws in the idealized standard deviation, while it is included in the latter. Therefore, it is no surprise that the point system for (1,0.5,0) is equal to the point percentage NSQF ratio for both the (1,0.5,0) and (2,1,0) systems, since they are all calculated without draws in the idealized standard deviation. When it comes to the point percentage, Fort emphasises that he follows note number four in Cain and Haddock, suggesting $0.5/\sqrt{m}$ as the idealized standard deviation. This is the standard deviation without taking into account drawn matches (equal to the description of the idealized standard deviation for a win-loss outcome league in Quirk and Fort, 1992). If, for example, Fort had used the same system when it comes to draws in the idealized standard deviation for both absolute points and point percent, the plots in his Figure 2 would have been similar for all seasons.

Table 1 in Fort (2007) has the following calculations:

$$NSQF(1,0.5,0) = \frac{ASD(1,0.5,0)}{ISD(1,0)}$$

which is equal to:

$$NSQF(100\%,50\%,0\%) = \frac{ASD(100\%,50\%,0\%)}{ISD(100\%,0\%)}$$

both for point percent systems for (1,0.5,0) and (2,1,0), while these are different from:

$$NSQF(2,1,0) = \frac{ASD(2,1,0)}{ISD(2,1,0)}$$

as is also the reason for the difference between absolute point and point percentage on the (2,1,0) score system in Table 1 in Cain and Haddock. The same explanation can be given for the following difference between the absolute point and point percentage score systems in Cain and Haddock:

$$NSQF(100\%,33\tfrac{1}{3}\%,0\%) = \frac{ASD(100\%,33\tfrac{1}{3}\%,0\%)}{ISD(100\%,0\%)}$$

$$NSQF(3,1,0) = \frac{ASD(3,1,0)}{ISD(3,1,0)}$$

The explanation that absolute point score and point percent NSQF are equal for a given score system can be done by following in Fort (2007), $\sqrt{C \times m} = \sqrt{C} \times \sqrt{m}$, where C = variance to the "unity structure" (own interpretation) of the league. The latter can for example be (1,0.5,0) and (1,1/3,0). $\sqrt{C}$ is therefore equal, both for ISD(points) and ISD(percents). Since the point score system is a scale H of a given unity system, the following relationship can be shown:

ISD(points) = $\sqrt{H^2 \times C \times m} = H \times \sqrt{C} \times \sqrt{m}$

ISD(points) = H x m x ISD(percent), where ISD(percent) = $\sqrt{C} / \sqrt{m}$

Since ASD(points) = H x m x ASD(percent) [same scale]:

NSQF(points) = NSQF(percent)

In other words, comparing a (2,1,0) system with a (1,0.5,0) system, either in points or in percentage, given that draws are included in the idealized standard deviation, gives the same NSQF ratio, and is therefore not related to any scale problem when it comes to measuring win dispersion in leagues. However, there are still problems to take into consideration. First. there might be a scope problem related to whether or not a draw should be included into the idealized standard deviation. as discussed earlier in this section (where I argue for *not* including it). Second, there are scale problems when comparing the (3,1,0) system with the other two other systems, because of the relative weaker valuation of draws. Note that NSQF (3,1,0) = NSQF (1,1/3,0) points = NSQF (100%,33 1/3%,0%), but the level of the ratio will here too be affected by inclusion of draws in the idealized standard deviation.

Multiple Outcome and Implications on Other Competitive Balance Measures

Some of the other measures shown in Chapter 2 are originally designed on basis of point score (from leagues that allow for tied matches). This is the case for:

- Gerrard's (1998) concentration ratio
- Michie and Oughton's (2004) five club concentration ratio (C5) and H Index of Competitive Balance (HICB)
- Koning's (2000) concentration ratio
- Seasonal Gini coefficients based on the cumulative number of points by the teams and the cumulative proportion of teams (Gerrard. 1998)
- End of season points for the winner (Gerrard, 1998)
- End-of-season point percent difference between the winner and the runners up (Gerrard, 1998)[92]
- Haugen's (2008) competitive balance measure, where maximal competitive balance is given by each match ending as a draw
- Groot's (2008) surprise index

Instead of using win percentage, Szymanski and Kuypers (1999) apply a measure for seasonal uncertainty that is based on the league's *end-of-season points' average standard deviation*. Changes in the number of teams and changes in the point system can be taken into account by adjusting seasonal mean points received by the teams, i.e. the *coefficient of variation*. Koning (2000) also measures win dispersion by standard deviation of points. Further. Koning applies a model that takes into account changes in home advantages over time and is invariant to changes in the point score system. as well

[92] Demmert (1973) also uses the differences between the first and the second place (number of games). but in a different manner (as an absolute quality measure for the single teams).

as it allows for draws. Competitive balance is calculated by estimating the standard deviation of teams' quality.[93]

3.4. Post-Season Tournaments

In the context of competitive balance, the tournament structure for deciding the championship winner is relevant. This can be explained by a simple example. Given a league with only one prize – the championship, performance incentives and the competitiveness within the league will be dependent on how this prize is won. Assuming that the best team is much better than the others, the interest for the league would be expected to be diminished, as described in the Louis-Schmelling paradox (Neale, 1964) and the Yankee-paradox (Vrooman, 1996). What would happen if this league ("one group league") introduced a championship playoff for the top four?[94] First, the focus would be moved from the first to the top four places (although the effect of this change might be relatively small if the point scoring system in the playoff incorporates the points achieved in the regular season), with a special focus on the fourth place. Second, uncertainty as to which team will win the championship can be significantly increased if the championship is won after a knock-out system in the playoff.

Real-world leagues are in general not unitary in structure, with some form of dependency between the regular season and other tournament(s). Generally, two main structures can be found. First, the regular season performance is the basis for qualification for a domestic post-seasonal playoff. This is the usual system for the NAML and for a number of European football leagues, as well as rugby league and rugby union in the UK and Australia. The second category of post-seasonal tournaments is related to a combination system, where *domestic* performance determines qualification for *international* team tournaments. In European football, this occurs through the system of tournaments organised by the UEFA, but similar systems also can be found in the other continents as well

[93] It is based on a statistical model to analyse results in football. More about modelling match results can be found in Dobson and Goddard (2001). They also include literature review on this topic. Groot (2008) suggests home advantage to be a small, but positive contributor to improved competitive balance. Home advantage in the context of competitive balance is beyond the scope of this thesis.
[94] See also Szymanski and Valletti (2005a) for effects of introducing a "second prize".

(en.wikipedia.org/wiki/List_of_football_%28soccer%29_clubs), such as the South-American CONMEBOL Copa Libertadores and the African CAF Champions League.[95]

The number of competitions for the teams in European football leagues is therefore, in general, higher than for the teams in the NAML. Hoehn and Szymanski (1999) use the term multiple leagues to explain that European football has a duality at team level because they are involved at both domestic and international levels of competition (where league tournament in season t is played simultaneously with the "post-seasonal" UEFA tournaments, which are based on qualification from the domestic league and cup tournaments in season t-1).

The UEFA competitions are, in general, highly valued by teams, both in sporting and financial terms. This is especially true for the lucrative UEFA Champions League tournament. This is a pan European team tournament arranged by the UEFA.[96] Although teams in European football leagues also qualify for the UEFA Cup and the UEFA Intertoto, the discussion below will be related to the UEFA Champions League, because it is the most significant of these tournaments, and that is has been more "structured" when it comes to qualification procedures. The qualification procedure for the UEFA Champions League is divided into a number of steps:

1. The number of teams that can qualify from a member league is determined by a seeding system.[97] The highest seeded leagues are allowed a maximum of four teams. At the time of writing, the top four teams in the English FA Premier League will qualify for next season's UEFA Champions League "system", while in Norway only the league champions will qualify.
2. The seeding system also determines at which round teams enter. In the case of the English FA Premier League, the top two teams qualify directly to the first group phase of the UEFA Champions League, while the teams finishing third and fourth

[95] The only international aspect by the NAML competitions, is, according to Andreff and Staudohar (2002), that NHL, MLB and the NBA includes teams cross-border (USA and Canada). However, Fort (2000) argues for similarities between the relationship domestic league and UEFA tournaments with the structure of the college sports in USA (see p. 436).
[96] According to Solberg (2004), a high share of the TV revenues goes to the teams. UEFA only keeps 25 percent of these revenues compared to 90 percent of the TV revenues from the UEFA European championship for national teams. A reason for this difference is the external threat to create a European Super League. The Italian Media partner had these plans at the end of the 1990s (see, for example, Hoehn and Szymanski, 1999).
[97] It is depending on teams' performance in the UEFA tournaments. For example, only leagues ranked as number 15 or higher have the possibility to have two teams in the UEFA Champions League (p. 39 in the Regulations of the UEFA Champions League, 2007/08, see www.uefa.com/newsfiles/19071.pdf).

enter the third qualifying round, the final qualifying round prior to the first group phase. The league champions in Norway currently enter in the second qualifying round.

3. The UEFA member league's seeding system is announced in advance so that teams in for example the English FA Premier League know at the start of the domestic league season how well they require to perform to qualify for the UEFA Champions League.

Entry into post-season tournaments represents prizes for league performance. Hence, prize concentration measures in these leagues should be extended to cover not only championship winners, but also qualification for post-season tournaments. For the NAML, this means including measure(s) of the concentration of playoff appearance for teams, in addition to the traditional measures of performance persistence, win dispersion and concentration of overall championship winners. Hadley et al. (2005) apply a Markov process model to measure the relationship between across season competitive balance and post seasonal playoffs in the MLB, by comparing turnover rates prior to and after the MLB strike in 1994.

In European football, a key question related to post-season play in the context of competitive balance is whether the UEFA Champions League affects the sporting quality balance among the different (domestic) member leagues. As relegation gives incentives to perform better (and hence incentives to invest in more playing quality) because of the fear of sporting and financial disadvantages, parallel incentives will exist bettering the upper part of the league. All teams qualifying for the UEFA Champions League tournament experience a significant increase in revenue compared to the other teams in their league. The effects on competitive balance can be expected to be the opposite of those of relegation. Ceteris paribus, stronger incentives to invest in playing talent for the leading teams will tend to weaken competitive balance. If for example one or two of the highest drawing power teams from a domestic league qualify season after season, one would expect their financial dominance to be higher than without the UEFA Champions League, and intuitively one would expect the "domestic" league's competitive balance to be negatively affected by this tournament. Lucrative financial and high sporting prestige post-season tournaments might also attract external investors to high drawing power teams, and further widen the distribution of financial resources and the gap between the "attractive" teams and the other teams.[98]

The examples above are based on the strongest teams increasing the quality of their team, not only to compete for the domestic championship, but also to increase the probability of qualification for financially lucrative post-season tournaments, such as the UEFA Champions League. This has been analysed by Hoehn and Szymanski (1999). The incentive for the most successful teams in domestic leagues to qualify for European level tournaments concentrates the distribution of playing talent within those leagues. Another issue, as emphasised in Barros et al. (2002), is that promoting competitive balance in one competition (e.g. the domestic league) might weaken the competitive balance in other competitions (e.g. UEFA Champions League).

Another interesting aspect, going beyond the original discussion in Sloane (1971), is the consequences for the regulation policy of domestic league of any concern with the international competitiveness of the best teams in the league. If one league has more restrictive rules than others, its teams might be less competitive than competitors in pan European tournament, ceteris paribus. This effect is likely to be much more important now in the post-Bosman era than at the time Sloane wrote the article, when there was a much more restrictive labour market in European football with limits on both within-league free movement of players out of contract as well as restrictions on player movements between leagues.

So although Downward and Dawson (2000) claim that the uncertainty of outcome hypothesis provides clear support for cross-subsidisation policies, such actions are more problematic in European football, because the redistribution of resources will impact on the competitiveness of the stronger teams in both domestic and European tournaments, and it is possible that the uncertainty of outcome effects may run in different directions. This is also suggested in Noll (2002), who concludes that the "arms race" effect (especially in the UEFA Champions League) might also obstruct domestic leagues in effectively introducing salary cap. This kind of regulation might reduce the playing quality for the best teams, since the labour market in the EU (including associated

[98] Perhaps this is what can be seen during the last seasons in English football, but it can also be related to a possibility to change the rules for distribution of media revenues to maybe be more performance based, or even more extreme: to individual deals. Such situations might increase the difference between high drawing power teams and other teams in the league, and hence further worsen the competitive balance.

members) is open. Further, Sandy et al. (2004) emphasise that the invariance proposition will not be valid if the effects of a valuable playoff are introduced (p. 98):

> If the chance to be in the limelight of post-season play was more valuable to the fans in the small market team because it was less common for them, introducing the post-season playoffs to the model would overturn the invariance results: that is, it would change the competitive balance. The model with post-season play is more complicated to diagram because it requires a discontinuity (i.e. a jump in the marginal revenue function due to reaching the threshold of wins for post-season play). To be realistic, the model would have to include uncertainty because the threshold would depend on winning percentage of many other teams.

These considerations by Sandy et al. suggest that the prize structure in complex league structure might affect competitive balance through incentive effects.

European football currently combines an integrated cross-European tournament structure (i.e. the UEFA Champions League) and an integrated international labour market. The financial situation in the domestic leagues is an important determinant for how well teams from a specific member league perform in the UEFA Champions League. This can create different incentives for different member leagues. At least two main questions arise:

1. Will the best teams in the different domestic leagues put pressure on the other teams, through the governing body, to protect their competitive advantages in order to maintain their international competitiveness?
2. Will the governing bodies want the bigger market teams to be successful in their league to increase the probability that these teams perform well in the Champions League? There are two reasons for governing bodies favouring the bigger market teams beyond any "political" pressure from these teams. First there is a financial incentive. The governing body receives "solidarity payments" from UEFA on basis of performance in the UEFA tournaments. In the 2002/03 season Norway received €659,315 (one team in the UEFA Champions League) and Scotland €475,913 (no teams in the UEFA Champions League) in solidarity payment, compared to €357,228 for Norway (no teams in the UEFA Champions League) and €1,156,167

for Scotland (2 teams in UEFA Champions League) the next season.[99] Second, members of the governing bodies may gain utility from the prestige of having team(s) from their league competing in the UEFA Champions League.

If governing bodies are concerned with ensuring the international competitiveness of their stronger teams, this can lead to domestic leagues preferring individual TV deals rather than collective deals that involve greater redistribution of income away from the stronger teams to the weaker teams (Solberg, 2004). It can also lead to other measures that favour the stronger teams domestically such as changes to the match schedule.[100]

The potential problem for a team budgeting with uncertain revenue from the UEFA Champions League is that revenues become more directly dependent on sporting performance. If such a team fails to qualify for a financially lucrative tournament, a significant share of the predicted revenues will be lost, causing possible financial difficulties. These financial problems can be temporary if the team qualifies for the tournament next time or the team has a cost policy where salary costs are performance-related to a significant degree. "Underperformance" over a number of seasons can lock teams into a vicious circle of decline experiencing a series of cumulative financial and sporting crises. Performance persistence in domestic leagues can be weakened at least in the short term if high revenue teams lose the financial advantage from qualifying for European tournaments.

On the other hand, a bigger league, where a number of teams secure playoff qualification at an early stage, might find that these teams do not perform as strongly as when playoff qualification remains undecided. Larsen et al. (2006) suggest that this might happen because teams rest some of their better players. They further hypothesise that teams with no chance of qualifying might use their last games to evaluate younger talented players. These factors might lead to reduced contest legitimacy (Gerrard, 2004b). However, leagues having within-league post-season playoffs usually have some

[99] Sources: www.uefa.com/uefa/Keytopics/kind=16384/newsId=80851.html, www.uefa.com/uefa/news/Kind=4096/newsId=219264.html and www.uefa.com/competitions/ucl/history/index.html.

[100] For example, in England, teams competing in the UEFA Champions League or the UEFA Cup are awarded a bye into the third round of the League Cup tournament. In addition, without having any suspicions, incitements might increase the probability of unethical biases when selecting teams for fair play prizes, because it can lead to qualifications for the UEFA Cup for better teams in a weak sporting season (see for example www.uefa.com/newsfiles/536640.pdf). This prize is based on both quantitative and qualitative variables (for Norway, see, for example, www.fotball.no/t2.aspx?p=58240&cat=51834).

incentives for already qualified teams to be fully competitive through to the end of the season. This can be the prize of winning the regular season tournament or divisional tournaments, as well as making the playoff schedule dependent on regular season performance with better performance in the regular season rewarded by a better seeding in the playoffs and a more favourable playoff schedule. An example of advantage in playoff seeding can be found in the Australian National Rugby League in many seasons, especially until the middle of the 1990s. For example in 1973, when the top five teams (out of 12) from regular season qualified for the semi-finals, the structure was as follows:[101]

Minor Preliminary Semi: Rank 4 – Rank 5
Major Preliminary Semi: Rank 2 – Rank 3
Major Semi: Rank 1 – Winner Major Preliminary Semi
Minor Semi: Loser Major Preliminary Semi – Winner Minor Preliminary Semi
Preliminary Final: Loser Major Semi – Winner Minor Semi
Grand Final: Winner Major Semi – Winner Preliminary Final

Even if the definition of competitive balance given by the Blue Ribbon panel on Baseball Economics (Levin et al., 2000) was argued to be difficult in Chapter 2, it is very interesting in the context of post-season tournaments, since post season play is used as benchmark. This means that in a league consisting of post-season playoff, perfect competitive balance should give all teams the same probability to qualify. The Report emphasises the continued domination of the playoffs by few big market teams, and that this represents a significant threat to the economic future of the MLB.

Post-season tournaments can also be held for poorly performing teams, such as relegation playoffs in some European football leagues. Relegation procedures can affect incentives, and this in turn impacts on win dispersion. If the bottom three teams are automatically relegated in a given top division, with a big gap between the fourth and third weakest teams, they would not have "anything to play for" in the final part of the season. If the same top division had a system where the third weakest team went to a relegation playoff, the teams at the bottom would still have "something to play for", and, hence, incentives to perform well would be higher than in the case of an automatic relegation system. These incentives due to relegation playoff could improve win

[101] This is based on the following reference: stats.rleague.com/rl/seas/1973.html. This is different from the scheduling method in the Scottish Premier League for the five "playoff" matches, as is according to Lenten (2007, p. 2) based on: "....a pragmatic method to determine a "fair" distribution of home and away matches".

dispersion. If several of the weakest teams have to compete in a relegation playoff to retain their current divisional status, this would have similar incentive effects as a championship playoff for the stronger teams at the top of the league.

Measures

Results in the UEFA Champions League can be applied for reflecting competitive balance across leagues in Europe. For example, Gerrard (2004b) analyses the distribution of nations represented in the semi-finals of this tournament, and found that the share of Big Five (i.e. England, France, Germany, Italy and Spain) teams increased from 65 percent to 95.8 percent between the periods 1992/93 - 1996/97 and 1997/98 - 2001/02. In the latter period only one team from outside the Big Five managed to reach the semi-finals of the UEFA Champions League.

Hadley et al. (2005) develop a Markov process measure for playoff qualification. This "state-dependent model" (p. 382) has two states: winning (i.e. qualification) or losing (i.e. non-qualification). Persistence is defined as the same state occurring in two consecutive periods. Lewis et al. (2007) measure the expected relative frequency of playoff qualification if all teams have an equal probability of qualification. Lee (2004) focus on the difference in performance between the team in the last qualifying and the closest non-qualifying team.

Alternatively, the mean win ratio for the qualified teams can be compared with either the mean win ratio and/or the highest win ratio among the teams that do not qualify. But there are potential problems with these calculations as that they might be difficult to compare, both within and between leagues, because of variations in the number of teams in the leagues in different seasons, as well as changes in the number of teams that can qualify for the playoffs. Playoff procedures changed frequently during the period of analyses of these studies.

Because of structural changes in relation to playoffs in the NAML (see Lewis et al., 2007), measuring playoff concentration is not straightforward for these leagues. One way to compare between and within leagues over time is to follow the same procedure as the NSQF ratio, which is to compare with the expected distribution in an idealized situation. This procedure can be applied to the *HHI* as a measure for *the concentration* *of playoff qualifiers*. The idealized HHI is calculated using the binomial distribution with equal probability of qualification.

I suggest that the idealized HHI should be computed by calculating the probability of playoff qualification for each team in any given season. These probabilities are aggregated for each team for the given analysis period to show the expected number of playoff qualifications for each team given perfect competitive balance. The "actual HHI" for the given period is then computed. The actual HHI based on the distribution of playoff contestants over the period can be compared with the corresponding idealized HHI. This form of the HHI will be called the *GK PO-ratio* (Gerrard-Kringstad Playoff-ratio):

GK PO-ratio = Actual HHI/Idealized HHI

Because the HHI will put relatively more focus on the concentration of teams qualifying for the playoffs, an alternative measure is suggested. This is based on the root of the aggregated squared differences between the actual number of playoff places for a given team and its expected number of places over a given period of time under conditions of equal probability of playoff qualification:

$$PO\ deviation = \{\sum_{i=1}^{i=N} (\text{Actual number of playoffs}_i - \text{Idealized number}_i)^2\}^{0.5}/\text{N}$$

where i = team number 1, 2,....., N. This alternative measure has relatively more weighting on the non-qualifiers, and this weighting is higher the more playoff places that are available. Both measures are related to differences between actual playoff and the idealized qualifications.

3.5. Unbalanced Match Schedule

One of the features of the simple league structure in Chapter 2 is a pure round-robin tournament representing a perfectly balanced match schedule. This has been the system used in England since the formation of the Football League in 1888. Each team plays against each other the same number of times, home and away.[102] This is the norm in many European football leagues. but there are a number of exceptions. For example, in the current system in the Scottish Premier League, each team plays every other team three times during the regular season play.[103] Perfectly balanced match schedules are not norm in the NAML. It is only in the NBA that all teams meet each other during a season. Noll (2003, p. 532) describes an unbalanced schedule as a league where "each team plays some teams more than others". On this basis, all NAML have unbalanced match schedules. In the context of competitive balance, the NFL's approach of making the current season's match schedule dependent in part on performance in the previous season (with better performing teams being matched up more frequently with other strongly performing teams) is particularly noteworthy.[104] Lenten (2007) claims that the match schedule in the NHL and MLB is unbalanced with teams being scheduled to play more frequently against "local rivals" (p. 20). Lenten (2007) also notes that teams' total travelling distance is also a determinant in the match schedule. There are also unbalanced match schedules in Australia in both the AFL (i.e. Australian Rules Football) and the NRL (i.e. rugby league). According to Lenten (p. 6): "...the unbalanced schedule is somewhat a bone of contention to many fans, especially since the ensuring of high-drawing matches being played more often is given priority over team-quality equalisation in designing the fixture. Nevertheless, these leagues persist with the unbalanced schedule, since they are happy with the current number of teams and season length". Lenten (2007) comments that the unbalanced schedule is nothing new, and was common for many years in English country cricket.

In European football, some leagues are arranged such that each team plays other teams three times during the season. This can be called a partial round-robin system, where the number of home and away matches between two particular teams is unequal. In the competitive balance context, the key issue is whether the distribution of home field advantage is random or allocated on the basis of team performance to improve the (expected) distribution of sporting outcome.

---

[102] Among the reasons to play this way might be to capture the well-known home field advantage. More about home field advantage can be found in for example Koning (2000), Dobson and Goddard (2001), and Forrest et al. (2005).

[103] Lenten (2007) defines the structure of the Scottish league different from this thesis, by defining it as unbalanced match schedule instead of post season playoff when it comes to the last part of the season (the five matches in either the "Championship" League or in the "Relegation" League) (see p. 2 and 3 in Lenten).

[104] Maxcy and Mondello (2006, p. 347) define "imbalanced schedule" as a situation where "teams with similar records from the previous season play each other".

Comparing the definitions in Chapter 1 of competitive balance and sporting quality balance, it is clear that if an unbalanced match schedule in a tournament can affect the relationship between the distribution of sporting quality and sporting outcomes, the distribution of sporting outcomes is more appropriate as a general definition of competitive balance, with the distribution of sporting quality and the match schedule as two of the causal drivers of competitive balance.

If the tournament match schedule provides an advantage for the weakest teams, and hence a disadvantage for better teams, one would expect the distribution of outcomes (e.g. end of season table) to be narrower than the distribution of sporting quality. Unbalanced match schedules can be seen as akin to a handicap system in horse racing or golf.

Another issue of relevance for competitive balance is that the revenue distribution might be affected by (an unbalanced) match schedule. For the better teams, both opponents of higher sporting quality and improved match uncertainty will increase demand, although eventually reduced chances of obtaining sporting success might weaken the demand. For the weaker teams, both increased probabilities of sporting success and increased match uncertainty will tend to be positively related to increased demand, but on the other hand, weaker opponents will tend to be negatively correlated with attendance. It is possible that the demand effects of unbalanced match schedules may actually increase revenues for the stronger teams but have little overall impact on the revenues of weaker teams. Thus an unbalanced match schedule might only improve competitive balance in the short run and perversely actually worsen it in the longer run through the financial benefits to the stronger teams.[105]

Suppose that an unbalanced schedule is introduced to improve the competitive balance in a league, and that the championship winner is decided on the basis of a championship playoff. This means that the highest ranked teams at the end of the regular season qualify for the championship playoff. Ceteris paribus, the unbalanced schedule should now increase the possibility for weaker teams to qualify for the championship playoff. However, given profit maximizing behaviour by the stronger drawing teams, one would

[105] According to Lenten (2007), Paul (2003) and Paul et al. (2004) are analysing effects on attendance from unbalanced match schedule in NHL and MLB, but these effects might be different from the analyses above because these leagues are unbalanced with regards to local rivalry, and not from sporting quality.

expect that these teams would invest more in playing talent to secure playoff qualifications (given that the playoff is profitable, and that incentives will increase as the expected payoff rises). For simplicity, assume that the strongest teams would prefer no change in the degree of competitive balance in the regular season (in order to maintain the same probability of qualification for the lucrative playoff), so that it would be expected that the sporting quality of the stronger teams would increase compared to the weaker teams. The conclusion from this simple analysis would be that the win dispersion in the regular season would continue to be approximately the same, while the sporting quality balance would be worsened with no improvement in the probability of weaker teams qualifying for the playoff. The concentration of championship winners could be even stronger. In other words, the "total competitive balance" (that takes into account both within and across seasons) might be worsened (as will the profit among the strongest teams, ceteris paribus). This can be called the "unbalanced match schedule paradox".[106]

Analysing the effects of unbalanced schedule on competitive balance is beyond the scope of this thesis, but intuitively the expectations are that there is likely to be some impact in a league where sporting quality differs between the teams. In a league where all teams are of equal sporting quality, the match schedule should not affect competitive balance, provided either that there is no home field advantage, or that teams play the same number of home and away matches if home field advantage exists.

Implications for Competitive Balance Measures

Defining competitive balance by sporting outcome implies that competitive balance measures such as the NSQF ratio do not need to be adjusted for an unbalanced schedule. Instead an unbalanced schedule can be treated as a possible explanatory factor for differences in competitive balance between leagues and/or over time. [107]

An *extension of the Goossens index* might be a possible measure to better show the relationship between ex ante quality balance and ex post comparable win dispersion measure. This is because this measure uses the perfect competitive dominance as benchmark, and can be applied in a situation where teams are ranked ex ante and

[106] Depending on the exact design of the match schedule.
[107] See Lenten (2007) for further references to a number of studies concerning these issues.

adjustments are done with regards to the given match schedule. This is exactly what Utt and Fort (2002) suggest, including unbalanced match scheduling and interdivisional and interleague play, following Fort and Quirk (1997), for (within seasonal) competitive balance comparison using the Gini coefficient. Utt and Fort (2002, p. 372) consider this to be very complex to apply to the MLB.

3.6. Prizes in Professional Team Sports

Szymanski (2003c) argues that prizes are widely accepted among both economists and the general public as an incentive driver in individual sports, but are seldom considered as such in team sports. However this is only the case if prizes are defined only as direct financial rewards based on performance.[108] In the case of professional team sports, it is more appropriate to widen the notion of prizes to include sporting prizes such as championship titles, playoff qualification, promotion and retention of divisional status (i.e. avoidance of relegation) (see Gerrard, 2006b), as well as financial prizes. The specific structure of sporting prizes is a key characteristic of a complex league with important implications for competitive balance. The prize structure has two main impacts on competitive balance: (direct) demand effects and incentive effects. The demand effects relate to the uncertainty of outcome hypothesis and match significance. Demand is not only driven by the closeness of individual matches, but also the closeness of the competition for the various tournament prizes. The demand effects depend on the fan motivation, and are hence likely to be related only to sporting prizes, whereas the incentive effects for team owners, coaches and players are likely to depend on both sporting and financial prizes.

The first step in this analysis is to identify the prizes in league tournaments. This topic has been largely ignored in the literature on team sports. The analysis of prizes in sports has largely been in the context of individual sports such as tennis and golf, where the prizes are largely financial apart from the sporting prize of winning the tournament. "The golf economy is one of the few examples of a pure prize economy" (Scully, 2002, p. 235). Financial prizes directly linked to tournament performance in many individual

[108] It looks like Szymanski in various articles (Szymanski, 2002, 2003b, 2003c) focuses on the direct financial part of prizes, and maybe therefore uses the term promotion when a team promotes from the second to first level, and the term relegation (or penalty - Szymanski, 2003c; Szymanski and Valletti, 2005b) is used for relegating teams (as of course are the usual terms) instead of using prizes as terms. However, the term "prize of first draft pick" is used in Szymanski (2003c, p. 61). This thesis will categorise all these issues as prizes.

sports are the principal income source for the competitors. In contrast, the prize structure in many team sports is much more complex and much less financially driven. This is highlighted by Szymanski (2003b, p. 1169):

> While most individualistic sports offer substantial financial prizes to the winners, this is usually not the case in (professional) team sports. The team that wins a league championship may receive a cup, and team members may receive substantial bonuses, but the owners of the team in general stand to gain little or no direct monetary gain (i.e. prize money) from winning a championship.

Further, Szymanski (2003b) emphasises that "participation in the playoff or final stage can be extremely valuable, and also that sponsorship income and merchandising are likely to be substantially increased by winning a championship, and that these factors will impact on decision making in much the same way as an explicit prize" (p. 1169). In a footnote, Szymanski also emphasises that (p. 1169): "Unlike a prize, the value of merchandising and related opportunities tends to differ between teams (e.g. because market sizes differ) and hence this kind of incentive promotes asymmetry". This incentive asymmetry is relevant in the context of competitive balance, because if there are prizes that are especially valuable for given teams, these teams will have an incentive to increase investment in playing talent compared to other teams.

Extending the concept of the prize structure in professional team sports leagues can be done by following the suggestion of Szymanski to include different rewards, sporting and financially that teams can achieve by sporting performance.

The most general difference in the prize structure between leagues is the difference between the open and closed leagues. In European football, characterised by a promotion and relegation system, one of the most significant prizes for weaker teams is avoiding relegation and retaining the team's current status in the merit hierarchy. Currently, only one of the top divisions among the members in the UEFA is closed, the San Marino league.[109] In contrast, the NAML are all closed leagues, where participation is not dependent on (recent) sporting performance. In closed leagues, such as the

[109] From the 1996/97 season the league has been closed (see, for example, en.wikipedia.org/wiki/Campionato_Sammarinese_di_Calcio and www.rsssf.com).

NAML and Australian Rules Football, there are prizes for weaker teams through the drafting procedures.

The "rookie drafting" is a system that regulates the recruitment of new players into a league. According to Andreff and Staudohar (2002, p. 39) it is "a characteristic of the American system", but it is also used in other leagues as well, such as in Australia. Rottenberg (1956) describes the "draft or selection rule" as a system that gives players in lower classification leagues the possibility of moving to a higher classification league. In general, the reverse-order-of-finish draft system allocates playing talent such that the worst-performing team in the previous season will be the first to pick a new player in each round of the draft (see, for example, Eckard, 2001; Fort, 2006a). It was introduced in the NFL in 1936 with the intention of improving competitive balance (see, for example, Fort and Quirk, 1995). The other NAML adopted the draft system later.[110]

Since the team that picks the first player will have an advantage compared to the other teams, the draft system can be considered as a "sporting prize".[111] However, the draft system is a sporting reward but not one that impacts on immediate sporting performance directly (unlike relegation), although a favourable position in the draft may yield positive sporting (and/or financial) benefits in the future, as described in Grier and Tollison (1994).

There are different ways of organizing a draft system. For example, a weighted lottery can be used to decide the order in which teams make their draft picks. Analysing the NBA and the use of different systems for drafting, Taylor and Trogdon (2002) find empirical support for tournament theory that a system of a "pure" reverse-to-order draft gives incentives for teams out of play-off contention to underperform at the end of the season in order to obtain a better draft position. This would also be the case with a weighted (as opposed to a uniform) lottery system for drafting.[112] The design of a draft system can affect the incentives for winning among the weaker teams in a league. In a perfect reverse order drafting system, there are incentives for teams to perform poorly. These perverse incentives are inconsistent with contest legitimacy (Gerrard, 2004b). In the competitive balance context, these incentives might actually lead to a worsening of

[110] NBA in 1949, MLB in 1965, and NHL shortly after (see, for example, Quirk and Fort, 1992).
[111] This is done despite that Rottenberg (1956, p. 249) claims the advantage to be "largely illusory".
[112] For example has the NHL (in 2005) an "Entry Draft" procedure that is based on a weighted lottery system (www.nhl.com).

ex post competitive balance since underperformance among the weaker teams will widen the win dispersion.

An interesting alternative possibility for organising the draft system is to have some form of "draft playoff" in which the weaker teams out of contention for the post-season playoff qualification would compete to win a better position in the draft picks. This would increase the incentives for these teams to be the best among the weakest, and should reduce the incentive differences between weaker teams in open and closed leagues. The significance of the reduction would be a function of the strength of the prize difference between the leagues. The prize for not being relegated would always tend to be much greater than any of the alternative prizes available to the weaker teams in a closed league. Hence, ceteris paribus, an open league will generate better competitive balance than a closed league.[113] The impact of relegation in open leagues and draft systems in closed leagues provide two clear examples of how competitive balance can be affected by incentive effects of the league's prize structure.

A review of studies in Szymanski (2003b) measuring the effects on competitive balance after introduction of a draft system shows mixed results with cases of both insignificant and significant improvement.[114] These empirical results indicate that the positive long term effects on competitive balance more than offset any negative incentive effects of reverse-of-order draft incentive in regular season play.

The other main difference in league structure is whether prizes are decided on the basis of sporting performance within the regular season or a post-season tournament (qualification for which is a prize in itself). A closed league with no playoffs has only one prize, the championship, while a league with a championship playoff has two prizes - qualification for the playoff, and winning the playoff (i.e. the championship).

One example of (direct) performance based financial prizes in European football is the performance based allocation of income from the collective selling of broadcasting rights in the FA Premier League in England. Szymanski (2003b) discusses how 25 percent of media revenues are based on performance, with the Premiership champions receiving twenty times as much as the team finishing in last place. According to for

[113] See Szymanski (2003b) for further discussion.
[114] These studies are Daly and Moore (1981), La Croix and Kawaura (1999), Fort and Quirk (1995), and Grier and Tollison (1994).

example Szymanski (2003c), it is based on the following formula for a 20 teams league: Broadcasting prize = f[(21-r)/210], where r is the rank in the final standing.[115] Elsewhere in European football there are other examples of performance based financial prizes such as in the UEFA Champions League and the Scandinavian Royal League[116].[117]

Based on the discussions above, prizes in the context of professional team sports can be divided into two types. *Financial prizes* are direct financial rewards based on sporting performance, while *sporting prizes* are prizes awarded on the basis of sporting performance, and are not "directly" financial. Sporting prizes in professional team sports generally consist of the championship, merit qualifications (promotion, qualifications to post-seasonal tournaments) and avoiding relegation.[118] In addition, more favourable future playing schedules and drafting rights can also be regarded as sporting prizes. Financial and sporting prizes are not independent, since they are both based on sporting performance. They may also have financial consequences through the impact on future sporting performance, but are not necessarily overlapping each other. Importantly, sporting prizes in team sports are hypothesised to be a demand driver, due to uncertainty of seasonal/tournament outcomes, as shown in Jennett (1984),[119] and hence have financial effects for the teams.

As argued above, prizes can affect incentives for sporting performance, and hence either ex ante competitive balance and/or ex post competitive balance. At the ex ante level, prizes can affect incentives related to the relative quality of a team, if certain teams can expect to profit more than others from the prize structure. For example, the introduction

---

[115] More generally: (N+1–r) / [(N+1)*0.5N].

[116] See for example: int.royalleague.com/page.aspx?id=97 and www.uefa.com/newsfiles/19071.pdf.

[117] An interesting interview emphasises that performance based financial prizes can be a motivating factor for a team. The coach for the Norwegian team Odd Grenland claims in a newspaper interview (pub.tv2.no/nettavisen/sport/tippeliga/oddgrenland/article474072.ece) that to be number eight (out of fourteen) should be a motivating factor for the team at the end of the season because that will give the team a two week period of training at La Manga in Spain during the winter for free instead of only one. According to the coach, this should be very important because the team is the one with lowest revenue in the league (in the 2005 season).

[118] There are also places in these tournaments that will be won by performance in domestic cups. For example, the teams in the FA Premier League in England compete in two domestic tournaments in addition to the league. This is the League Cup and the FA Cup (prizes here are UEFA cup participation). In addition, financial prizes are also won by performance in these cup tournaments. See the following web pages: www.football-league.premiumtv.co.uk/staticFiles/4f/7a/0,,10794~96847,00.pdf, www.thefa.com/TheFACup/TheFACup/NewsAndFeatures/Postings/2007/06/FACupPrizeFund and www.thefa.com/TheFACup/TheFACup/NewsAndFeatures/FAC_Cardiff_UEFACup.htm.

[119] Demand studies show that match/championship significance is a significant variable (see, for example, Jennett, 1984; Downward and Dawson, 2000; Dobson and Goddard, 2001 for reviews).

of the financially valuable UEFA Champions League may have changed the investment policy of some (high drawing power) teams, compared to other teams, and hence contributed to a deterioration in the sporting quality balance in the league.

In summary, sporting performance is rewarded by a number of prizes in professional team sports. These prizes may have financial value, but the value can differ significantly, both with regards to the given team's attributes (such as drawing power), and with the sporting status of the prize. For example, there is only a relatively small financial difference between finishing first and second in the English FA Premier League, but a huge difference in prestige between these two outcomes. This relationship and the typical prizes in professional team sports can be summarised as follows:[120]

		Prestige Trophies		
		Merit Qualification		
Sporting Performance	→	Future Match Schedule	→	Financial Value
		Draft Picks		
		Financial Prizes		

In sports, prizes can be defined as tournament payoffs for a certain sporting performance. Prizes (sporting and financial) should be split into two types for professional team sports: graded and threshold prizes. Graded prizes can be defined as prizes where payoff is a direct function of sporting outcomes. One example is the performance based part of broadcasting revenues for teams where the deals are collectively negotiated, such as in the English FA Premier League. Future (season t+1) match schedule might also be categorised as graded if it is based on (the reverse of the) performance in this season (t).

Threshold prizes are rewarded on basis of reaching a certain limit (threshold) in sporting output. Merit qualifications, such as continuing (no relegation) in the division and qualification for post-season play, are examples of threshold prizes. Winning the championship (prestige trophies) is another example of this kind of prize.

Each prize category in this overview can be treated as graded and/or threshold prizes. For example, the classification of draft picks is related to the system in the given league. A pure reverse-order system is graded, while a combination of reverse-order among the

[120] Direct financial effects from different prize and league structures will not be discussed in this chapter, even if closed leagues may have financial advantages for the teams due to incentive effects in a league with relegation (see, for example, Szymanski, 2003b; Szymanski and Ross, 2000).

upper half and uniformly weighted lottery among the bottom half combines the graded and threshold prize systems.

In general, all prizes (including sporting prizes) can be related to financial value. The financial value for a prize can be split into two parts: certain payment (direct financial value, such as broadcasting revenues) and future payment (indirect financial value, such as payoff from qualification to the UEFA Champions League, and negative future financial effects from relegation). Hence each team faces a prize payoff function in relation to performance:[121]

$$PPO_i(r) = f(r) + g_i(r)$$

where:

$PPO_i(r)$ = the prize value for team i related to its sporting outcome r.

f(r) = financial payoff from graded prizes, as a function of sporting outcome r.

$g_i(r)$ = financial payoff from threshold prizes for team i as a function of sporting outcome r. It is related to (present value of) future financial payoff from winning a threshold prize, such as qualification for UEFA Champions League.

Note that a significant difference between the NAML and European football systems appears when it comes to the course of $PPO_i(r)$. In European football all prizes are upward driven, with break points (avoid relegation, UEFA Champions League), so the better sporting outcome the higher value, both sporting and financially for a given team. This structure of prizes can be called monotonic, and differs from the NAML system, where prizes are both upward and downward driven. These leagues have this two-way non-monotonic prize system, because additional to the upward prizes, both draft pricks and future schedule might be positive related to *weaker* sporting performance. Another difference between these two continents is that a given team in European football can only win one domestic prestige trophy within a league tournament - the championship. In NAML the tournament structure makes it possible for a given team to compete for

[121] Further, the difficulties of prizes in teams sports are emphasised in Szymanski (2003b, p. 1169): "One might hope to see future research attempt to quantify the value of prize like elements in the different team sports."

more prestige trophies, such as being a division winner and a conference winner in the context of the same tournament.[122]

Prize concentration in perfect competitive balance can be related to a distribution of prize winners with the following expected outcome for each team:

$$E(team_i) = (z*t)/N$$

where:

$E(team_i)$ = expected number of times that the i^{th} team will win a specific prize.

z = the number of winners of a specific prize in any given season (z = 1 for the championship but z > 1 for post-season tournament qualification).

t = the number of seasons.

N = number of teams in the league.

In addition to having the possibility for measuring success-failure ratios on every single prize in a league, the existence of both positive and negative valued prizes in for example European football leagues gives an opportunity to create a new prize driven measure based on the difference between teams that are successful (win positive prizes) and teams that fail (are punished by negative prizes). By successful is meant a team winning a positive sporting prize, such as qualification for a European tournament, while relegation represents a failure prize. In addition to success and failure prizes, European football leagues have what could be called neutral prizes, which is to retain the current status. If, for example, the difference between success and failure prizes is small, this can indicate a high level of competitiveness. A measure can be called the "success prize - failure prize ratio":

$$\frac{(\text{Performance marginal success team} - \text{Performance marginal fail team})}{\text{Average league performance}}$$

or[123]

Win percentage marginal success team – Win percentage marginal fail team

[122] European teams can also win multiple prestige trophies in a single season, but by competing in multiple separate tournaments.

[123] Adjustment is not necessary because the average is 0.5.

3.7. Competitive Intensity

The current literature on competitiveness in professional team sports has not yet fully recognised the significance of real-world multi-prize complex tournaments. Zimbalist (2003a), Quirk and Fort (1992), Bennett and Fizel (1995), and Gerrard (2004b) describe competitive balance as a complex concept. Both Bennett and Fizel (1995) and Quirk and Fort (1992) emphasise that the competitive balance concept "refers to a number of aspects of competition". In Chapter 2 this complexity was considered in the context of a simple league structure in regard to time (ex ante and ex post) and the three dimensions (win dispersion, performance persistence and prize concentration). The current chapter has extended the consideration of the complexity of competitive balance arising from the complexity of tournament and prize structures.

Including multiple prize structures into the term competitive balance will make it even more complex than the multi dimensional focus in Chapter 2, and perhaps too complex to be operational. Hence, it might be appropriate to use a straightforward definition of the competitive balance concept that reflects the consensus that competitive balance refers to the distribution of the sporting outcomes among the teams in a league/tournament. Note that this consensus view does not preclude using the distribution of prizes as indicators of competitive balance. But I suggest that we should treat competitiveness relative to the league's prize structure separately from the competitive balance concept. Typically in European football, as, for example, in the English FA Premier League, the multiple prize system creates a series of "sub-tournaments". In the NAML there are several aspects relating to tournament structure, such as qualification procedures for playoffs and the differences in league structures. These aspects can all be included in a new concept called *competitive intensity*, defined as *the degree of competition within the whole prize structure of the league/tournament.*

The differences between the concepts of competitive balance and competitive intensity can be shown using the English FA Premier League as an example. Competitive balance reflects the differences in sporting outcome among all teams in the Premier League, while competitive intensity will give a picture of the intensity of competitiveness for the different prizes (or sub-tournaments) within the league. For European football this means that competitive intensity is a concept that tries to capture a league's overall competitiveness with regards to championship, direct entry to the UEFA Champions League, entry to UEFA Champions League qualifying rounds, qualification for the UEFA Cup, and avoiding relegation.

It is possible that the Premier League has a weaker competitive balance as a consequence of increased dispersion of sporting quality among the teams, but at the same time competitive intensity may have been improved, due to a changed prize structure, such as the increased number of teams qualifying for the UEFA Champions League.

In the FA Premier League, the measures of competitive balance in all three dimensions show weakened competitive balance over the last 20 years, but, contrary to the uncertainty of outcome hypothesis, fan interest measured by gate attendances has increased substantially. The NSQF ratio has increased from an average of 1.44 in the period 1986/87-1995/96 to an average of 1.61 in the next ten season period. The average Spearman's rank correlation coefficient has increased between the two periods from 0.44 to 0.62. In addition, the Herfindahl-Hirschman index on championship winner concentration has increased from 0.2 to 0.38. All these measures are consistent in indicating that competitive balance in the English top division has weakened during this period.[124] However, over the last 20 seasons, the average attendance has increased by more than 70 percent, even with admission prices having risen significantly. Of course, there are a number of factors affecting attendance, and for the English top division one of these may be greater exposure through TV coverage. The quality of the league has also risen significantly in the post-Bosman period with a significant influx of top foreign players.[125]

However, even if there are a number of external factors that have affected the interest for the league, other elements have changed as well. At the start of the period referred to above, teams from the English top division were excluded from the UEFA team tournaments; they are now fully involved again and a significant presence in the final knockout stages of the tournaments most seasons. In addition, structural changes in the UEFA Champions League and the UEFA Cup have tended to be advantageous for

[124] More detailed empirical results on competitive balance in the FA Premier League and other leagues will be presented in later chapters.
[125] The global football market, as described as one of the elements in Andreff and Stadudohar's (2002) MCMMG model, can be seen as a characteristic for the current FA Premier League.

teams from the biggest leagues, such as the FA Premier League. Previously, only one team (i.e. the league champions) could participate in the UEFA Champions Cup. Currently the top four teams from the English FA Premier League may now enter the UEFA Champions League. As a consequence, the structure of the English top division has changed significantly during these twenty seasons - from a time where two prizes were awarded; the championship and the prize to stay in the division, to the current prize structure where teams in addition to these two prizes can qualify for UEFA Champions League directly, win entry to the UEFA Champions League qualifications, qualify for the UEFA Cup and UEFA Intertoto cup. The competitive intensity has been increased due to the more complex prize system, and this may have offset the effects of the reduction in competitive balance on the uncertainty of outcome and the enjoyment/excitement of attending and watching games (see also Sandy et al., 2004).

Match significance in a simple league context is only related to the championship. In a complex league system, the number of matches of significance for prize outcome will usually be higher, ceteris paribus, because teams will also have matches of significance for playoffs and relegation. For example Cairns (1987) claims that because of restructuring the Scottish football league in 1975, more matches became significant ("counted", p. 262) in relation to the championship, qualifications for European tournaments and relegation. Because uncertainty of outcomes at seasonal level in complex league systems will be related to a number of prizes, match significance will depend on the prize structure. In other words, the number of prize significant matches is likely to be greater in a complex league system, ceteris paribus.

The way of interpreting uncertainty of outcome on seasonal/tournament level in Jennett (1984) has many similarities with the competitive intensity concept. Jennett claims that only focusing on the playing strength among the teams in a league might be too narrow to represent uncertainty at the seasonal/tournament level, because it is not capturing enough of the relationship between uncertainty and attendance, particularly the utility that spectators derive from matches that are significant for "the overall championship race" (p. 179). More specifically, the basic idea behind competitive intensity is related to Jennett's championship and relegation significance variables. Competitive intensity will capture these conditions, and reflect the number of prizes, their importance and the intensity around each prize.

Competitive balance and competitive intensity are not strictly independent concepts. Intuitively, competitive balance will affect competitive intensity because, ceteris paribus, a league with better competitive balance is expected to have more intense competition for the different prizes. However, this relationship is not unambiguous. If a league's sporting outcomes are concentrated around two poles, the top prizes and the relegation battle, the win dispersion will be wide but there will be a high level of competitive intensity. Since competitive intensity is strongly related to a tournament's prize structure, competitive balance is only one of the dependent variables driving the level of competitive intensity:

$$CI = f(prize_p, CB, u)$$

where:

$prize_p = prize_1, prize_2, \ldots, prize_p$; the different prizes should be weighted on basis of its relative importance.

CB = competitive balance, which is assumed to affect the competition among the prizes.

u = others variables affecting competitive intensity.

There are at least two advantages of introducing the concept of competitive intensity. First, it allows the competitive balance concept to be defined in a more limited and straightforward way. Second, it focuses more attention to the leagues' prize structure. The challenge remains to find appropriate aggregate measures for competitive intensity that take into account all prizes/outcomes of a league, the level of uncertainty of outcome for each, and degree of importance. The approach adopted here is a more disaggregated one, focusing on the measurement of the degree of competitive intensity for individual prizes.

In the NAML, the large number of divisions potentially increases the degree of competitive intensity, because a team's relative performance is dependent on both the teams in its own division, as well as the general standing in its conference. This might increase the level of match significance and competitive intensity related to winning the division, winning the conference, and qualifying for the playoffs. Competitive balance, on the other hand, is more independent of the tournament and prize structure, apart from any incentive and schedule effects, as discussed above.

From the perspective of competitive intensity, the unbalanced match schedule might affect the intensity through creating more matches of significance around the prizes, and possibly also more intense competition around the prizes, because more teams are fighting for the different prizes.

The three key elements in the competitive intensity equation above can be summarised by "the competitive balance circle", as shown in the following figure:

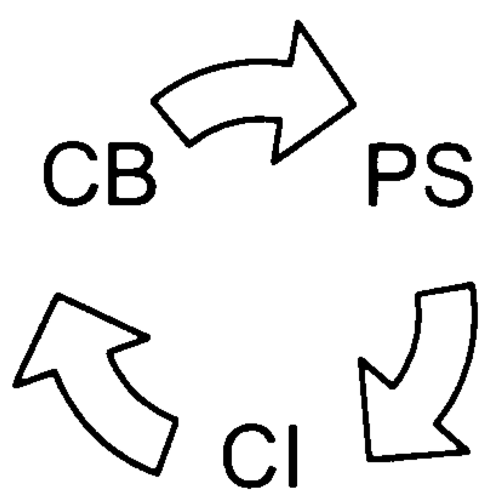

Figure 3.1: The competitive balance circle

The first element of the circle is related to competitive balance (CB) in a simple league context. The next element captures the multiple prize structure (PS) and the limitation of competitive balance as a means of reflecting the overall competitiveness of a tournament relative to all prizes, and hence the development of the concept of competitive intensity (CI) as the third element of the circle. But, as discussed above, the incentive effects created by the prize structure and the competitive intensity for each prize will in turn impact on competitive balance, taking the relationship full circle. These issues will now be explored empirically in the second half of this thesis.

4. Empirical Analyses of Competitive Balance across Leagues

4.1. Introduction

The most popular measure of competitive balance, the NSQF ratio, gives a picture of competitive balance related to a so-called idealized league, where all teams are assumed to have the same probability to win the tournament. Given the assumptions of the calculations of the idealized league, as shown in Chapter 2, an analyst will be able to calculate how far away a given league's competitive balance is from the ideal. Additionally, a league's governing body might also want information from other leagues for comparing competitive balance across leagues, as well as analysing how its "relative" competitive balance has changed over time.

This chapter's purpose is to empirically give a detailed picture of competitive balance in the football leagues in Europe, in addition to comparisons with other leagues. The current literature of competitive balance has mainly focused on competitive balance in the bigger leagues. For European football this means the top divisions for the Big Five leagues. The motivation for this chapter is to empirically measure competitive balance for *all* member leagues of the UEFA. Due to the great differences in such as the size between different European football leagues, they will be categorized in different groups in the presentation in this chapter. This categorisation of leagues makes it possible to extend analyses beyond the recent literature, and analyse whether the Big Five leagues are representative for competitive balance in European football, or if there are systematic differences between different groups of leagues.

The general picture of competitive balance will be related to the three dimensions described earlier. In other words, for each league, competitive balance is measured both with regards to win dispersion, performance persistence and prize concentration. This should take into account a criticism in Szymanski (2003b) that only focusing on win dispersion (standard deviation) will be too narrow, and will not take into account variation across seasons. The three basic dimensions will be measured by the NSQF ratio, the SRCC and the HHI on championship winners. These measures are among the most popular within their dimension.[126]

Prize concentration can be measured in two ways. First, the distribution of championship winners is a generally applicable, since all competitions have an overall winner. Second, in professional team sports, qualifying for playoffs can also be categorised as prizes, and playoff concentration is an alternative to championship concentration. For example, all the NAML have post-seasonal playoffs. In European football this part of the dimension is not straightforward. All UEFA member leagues have some kind of post-seasonal playoff, since performance in regular season can be used for domestic championship playoffs and/or qualification for the international UEFA team competitions (that can be seen as a post-season play). The last part of this chapter will apply complex league system measures related to multiple match outcomes, post-seasonal qualifications in some leagues, and promotion and relegation issues.

4.2. Literature Review

Empirical analyses of competitive balance have been part of the literature of the economics of professional team sports. This literature includes descriptive analyses of competitive balance in leagues, and in the beginning it focused on the NAML, both related to single league analyses and comparisons with other major leagues. Rottenberg (1956) counted the number of championships won by the best teams, and showed imbalance in both AL and NL. Quirk and Fort (1992) compare competitive balance within and between the NAML over time. Vrooman (1995) summarizes his findings of differences in competitive balance in the MLB, NBA and NFL (p. 985):

> Clearly the NFL is the most competitively balanced, MLB is slightly less competitively balanced, and the NBA is the least competitively balanced of the three leagues. These estimates of competitive balance.....are corroborated by every recent study....

More recently, similar analyses have been extended to other leagues as well. In Europe, the focus has been on (European) football leagues, both related to "single league trends"

---

126 This means that for example the Kendall's τ will not be used, even though it according to Groot (2008, p. 15) is "relatively easy to grasp how this measure is calculated".

over time, comparisons with other European leagues, as well as to NAML. Examples of empirical studies of competitive balance in European football can be found with Szymanski and Kuypers (1999), Hoehn and Szymanski (1999), Koning (2000), Szymanski (2001), Dobson and Goddard (2001), Szymanski and Smith (2002), Buzzacchi et al. (2003), Gerrard (1998, 2004b, 2006a), Bourg (2004), Dejonghe (2004), Cain and Haddock (2006), and Groot (2008). An interesting observation is found in Gerrard (2004b), showing that the Big Five leagues seem to have better competitive balance.[127]

When comparing competitive balance between NAML and a number of European football leagues, both Szymanski and Smith (2002) and Buzzacchi et al. (2003) emphasise that the competitive balance indications are dependent on how the concept is measured. Average within-seasonal measures, usually the NSQF ratio, typically find the open leagues in European football to be better balanced than the closed NAML, although Buzzacchi et al. find a trend towards similarity during the sample period. The latter can be explained through a negative trend on average for the European football leagues, while at the same time the NAML on average had a trend in the opposite direction, except from the 1970s. However, this is only partly confirmed in the calculations by the NSQF ratio in this thesis. For the Big Five leagues there are indications of a negative trend. For NAML there is only slight improvement, but the distance between the leagues is still significant (in favour of Big Five). Note that the comparison of competitive balance in three NAML (the MLB, NFL and NHL) with four European football leagues (Italy, Spain, Portugal and England) over thirty years in Szymanski and Smith (2002), find the NSQF ratio to be pretty similarly balanced on average on both sides of the Atlantic.[128]

Performance persistence, or dynamic competitive balance,[129] that is related to season to season changes in performance among the teams, is better in NAML measured by Buzzacchi et al. (2003) and Szymanski and Smith (2002). They conclude that the closed NAML are better balanced over time than the open European football leagues. Szymanski and Smith also emphasise that even if more teams are included in the top divisions in European football over a number of seasons, the across season competitive

[127] Four of the Big Five leagues are historically among the highly competitive group of leagues (when ranking 16 Western European leagues into three groups of competitive balance).

[128] However, the calculations in Szymanski and Smith (2002) differ from Buzzacchi et al. (2003). The latter is equal to what is used in similar calculations in this thesis.

[129] This term is for example used by Szymanski and Smith (2002) and Buzzacchi et al. (2003).

balance seems to be better in the closed NAML. The calculations in this thesis do not support this general conclusion, except from the last twenty seasons of the sample, where NAML is slightly better.

In general, Szymanski and Smith (2002, p. 122) describe the consequences of the structural differences between European football leagues and NAML by using the expressions "equality of opportunity" and "equality of outcomes". The first is related to that the European leagues are open for all teams, but are in general dominated by only a few. The second expression is related to the closed NAML that are not open to all, but where the teams that are included "have a fair chance of competing at the highest level."

Buzzacchi et al. (2003) give a preference of which of these two ways to measure competitive balance should be most weighted (p. 182): "We believe that the dynamic measure presents a better picture of competitive balance than the static measure". The differences between the NAML and European football leagues are characterised in Szymanski and Smith (2002, p. 109) as the difference between "the more cartelised North American leagues" and "the more 'competitive' European leagues".[130]

4.3. Data and Methods

Presenting empirical data on competitive balance within European football is not straightforward. By June 2006, the UEFA consisted of 52 member associations. These are:[131]

> Albania, Andorra, Armenia, Austria, Azerbaijan, Belgium, Belarus, Bosnia-Herzegovina, Bulgaria, Croatia, Cyprus, Czech Republic, Denmark, England, Estonia, F.Y.R Macedonia, Faroe Islands, Finland, France, Georgia, Germany, Greece, Hungary, Iceland, Israel, Italy, Kazakhstan, Latvia, Lichtenstein, Lithuania, Luxembourg, Malta, Moldova, Netherlands, Northern Ireland,

---

[130] See, for example, Hoehn and Szymanski (1999), Buzzacchi et al. (2003) and Szymanski (2004) for further description and overview over structural differences between the NAML and European football leagues.

[131] Later in 2006 Montenegro went out of the Serbia and Montenegro association. Montenegro created its own association that became a provisional member from the 5th of October 2006 (www.uefa.com/magazine/news/kind=16/newsid=469287.html), and became a full member in 2007 (www.uefa.com/footballeurope/countries/association=92853/index.html). Serbia continued as a separate association. Since these changes happened after the end of the sample period of this thesis, they do not affect the analyses.

Norway, Poland, Portugal, Republic of Ireland, Romania. Russia, San Marino, Scotland, Serbia and Montenegro, Slovakia, Slovenia, Spain, Sweden. Switzerland, Turkey, Ukraine and Wales.

All of these UEFA members have their own league, except Lichtenstein, where the teams play in the Swiss league.[132] In addition to the leagues given above, a number of former associations have been dissolved as a consequence of the restructuring at national level in Eastern Europe. These are the associations from Soviet Union, East Germany, West Germany, Czechoslovakia and Yugoslavia.

The number of European football leagues included in the sample are the "current" (2005/06) 51 leagues of member associations, plus the four "old" leagues in the Soviet Union, East Germany, Czechoslovakia and Yugoslavia. The total number of European football leagues covered during the sample period is therefore 55 leagues (Germany and West Germany are treated as one entity[133]).[134]

Two problems appear in analyses of comparisons of competitive balance over time across European football leagues. First, there is a significant group of leagues without UEFA membership during the whole sample period, particularly leagues in Eastern Europe. Second, leagues have undergone structural changes during the sample period, and some of these league structures do not fit with the measures used in this chapter. An example here is (West) Germany before the restructuring to the 1963/64 season, where the league was divided into five geographical divisions, with championship playoffs.[135] It is difficult to calculate definitive results on the NSQF ratio and the SRCC under this structure.[136] Therefore, complete comparisons of competitive balance between the UEFA member leagues cannot be made for the whole sample period.

The sample period in this part of the thesis is sixty seasons (the post Second World War period), from 1946(/47) to 2005(/06). It is therefore possible to combine different periods over time, such as six periods of ten seasons, four periods of 15 seasons and two

[132] There are also leagues where some teams participate in other leagues, such as that San Marino has one team in the (professional) Italian league (en.wikipedia.org/wiki/San_Marino_Calcio), and Wales has teams in the English league (www.football-league.premiumtv.co.uk/page/Home/0,,10794,00.html).
[133] This is following the UEFA (www.uefa.com/uefa/aboutuefa/matchassociations/association=47.html).
[134] There is also a number of European football leagues organised in associations outside the UEFA. These are not included in the sample of the thesis. These leagues can be found at www.rsssf.com .
[135] See for example www.f-archiv.de .
[136] In general, analyses of championship winners will be independent of structure.

periods of 30 seasons. In this chapter the choice is to concentrate on ten season periods, which might be seen as arbitrary. However, since this will make it easier to put the "new" leagues in Eastern Europe into position to be included in the last period, it makes sense not to use too long periods. In addition, the latter period can include consequences of the Bosman verdict and changes in the UEFA Champions League structure.

It is useful to "categorise" different European football leagues for the purposes of analysis. The Big Five leagues in European football is a generally accepted grouping. The analyses below will extend the categorisations of leagues on the basis of both location, date of formation, and financial size (as measured by average league attendances in the top division).[137] Additional to the European football leagues, the NAML are included for comparison purposes. The different groups are (additional to Big Five):

- 2WE = Netherlands, Turkey, Scotland, Portugal and Belgium
- 3WE = Norway, Sweden, Switzerland, Denmark, Austria and Greece
- EEW = Romania, Poland, Bulgaria, Hungary and Albania
- EEO = Soviet Union, Czechoslovakia, East Germany and Yugoslavia
- SOV = Soviet Union, Russia, Ukraine, Kazakhstan, Belarus, Georgia, Azerbaijan, Lithuania, Latvia, Estonia, Armenia and Moldova
- CZE = Czechoslovakia, Czech Republic and Slovakia
- YUG = Yugoslavia, Croatia, Serbia-M., Slovenia, Bosnia-Herzegovina and F.Y.R Macedonia. Bosnia-Herzegovina is only included in the championship winner concentration measure.
- GER = East Germany, West Germany and Germany
- MEL A = Andorra, San Marino, Israel, Northern Ireland, Malta, Faroe Islands, Cyprus, Finland, Ireland, Iceland, Luxemburg and Wales
- MEL B = Northern Ireland, Malta, Cyprus, Finland, Ireland, Iceland and Luxemburg
- NAML = NBA (National Basketball Association), MLB (Major League Baseball), NHL (National Hockey League), and the NFL (National Football League)

Analyses are not limited to the time after federations has joined the UEFA. This includes for example Israel and Faroe Islands.

---

[137] Source: www.european-football-statistics.co.uk/attn.htm.

According to Fort (2006a, 2007), it is important to calculate competitive balance for the leagues where the matches have been played. This is relevant for the MLB where the teams are divided into two leagues, the American League (AL) and the National League (NL). Until recently, the teams only played their matches within its own league (and no cross-league matches). However, the World Series (playoff) is based on both NL and AL.

The data on football leagues is collected from the rsssf.com, and from websites linked from rsssf.com. The analyses from European football leagues will only be on the top divisions. For the NAML, data is collected from their websites: nhl.com, nba.com, mlb.com and nfl.com, and from ESPN.com.

The empirical analyses will mainly focus on three issues. These are trends in competitive balance over the sample period within each group of leagues, comparisons of competitive balance between the different groups of leagues, and relationship between measures of the three basic dimensions. In addition, some of the leagues within groups will be highlighted. In other words, analysing a league's competitive balance requires a range of points to consider:

1. Cross-sectional comparisons to indicate the level of one league's competitive balance in comparison to other leagues.
2. Time series
    a. Comparing a specific league over time to indicate eventual trends over time in competitive balance. The statistical tests applied in this chapter are t-tests on basis of differences between two means (see, for example, Triola and Franklin, 1994).
    b. Across leagues. Comparing one league's trends over time with other leagues' trends.
3. Comparisons in the three dimensions of competitive balance.

Some calculations will be done for periods consisting of less than ten seasons. Because the use of average values of each league, adjustments/weightings should not be required for win dispersion and performance persistence. However, HHI is adjusted for "missed" seasons.

The main focus in this chapter is to provide a survey over competitive balance across and within leagues over time. The first part will concentrate on what can be seen as the three main measures in this thesis:

- End of season) win dispersion: NSQF ratio, where wins are given the value 1, and draw matches are treated as half wins. The idealized standard deviation is not adjusted for the possibility for draw matches.[138] In other words: [ASD(100%,50%,0%)]/[ISD(100%,0%)].
- Prize concentration: HHI of the championship winners.
- Performance persistence: SRCC (from season to season), based on the method to ex ante rank promoted teams (see Chapter 3).

The second part is related to more specific measures applied in complex league structures. This will only be done for a number of leagues.

4.4. Empirical Results

4.4.1. Win Dispersions

The NSQF ratio indicates the spread of sporting outcome among all teams in a league. A higher ratio reflects a relatively wider spread in the sporting outcome, and hence weaker competitive balance. As shown in Chapter 2, a ratio of one indicates perfect competitive balance in relation to what is defined as a simple league.

As described earlier in this chapter, win dispersion has not been calculated for a longer period in all European football leagues. This is taken into account in the different groups. For example in the Big Five group, Germany is not included in the win dispersion calculation in the first two periods. To avoid eventual distortion in comparisons over time, win dispersion excluding Germany is also calculated for all periods. The same principle is followed for the other groups. T-tests are done within each group from period to period. The null hypothesis is assumed to be no change in competitive balance, while the alternative hypothesis is that there is a change in

[138] See the discussions in Chapters 2 and 3.

competitive balance. In such a broad analysis, there are no prior beliefs on the direction of change in competitive balance,[139] so the tests are two-sided. In addition to the general comparisons over time and across leagues, there are also other general issues of special interest such as the impact of the introduction of the UEFA Champions Cup, expansion of UEFA Champions League, free agency (i.e. the Bosman ruling), and the split from bigger to smaller leagues in Eastern Europe. The first should be related to differences from period (I) to (II), while the other changes mentioned above are relevant for the difference between period (V) and (VI). For the restructured leagues in Eastern Europe, comparisons will be done between period (VI) and the previous period. The results on win dispersion are shown below:

Table 4.1: Ten-season win dispersion measured by the NSQF ratio in European football and NAML from 1946(/47) – 2005(/06)

	1946/47-1955/56 (I)	1956/57-1965/66 (II)	1966/67-1975/76 (III)	1976/77-1985/86 (IV)	1986/87-1995/96 (V)	1996/97-2005/06 (VI)	Sign I→VI
Big F. ex. Germ.	1.26	1.34	1.34	1.41	1.44	1.49	*
Big Five	1.26	1.34	1.32	1.42*	1.42	1.49	**
2WE ex. NL/Tur	1.49	1.69	1.73	1.78	1.82	1.81	
2WE	1.49	1.59	1.68	1.65	1.79	1.81	*
3WE ex. Greece	1.45	1.50	1.37	1.46	1.38	1.47	
3WE	1.45	1.50	1.45	1.47	1.41	1.56	
Norway	1.38	1.34	1.31	1.21	1.30	1.41	
EEW	1.57	1.32	1.30	1.21	1.41*	1.67*	
EEO	1.48	1.23*	1.19	1.29			
SOV[1]	1.57	1.49	1.29	1.25		2.23	
CZE[1]	1.38	1.18	1.00	1.16		1.55	
YUG[1]	1.59	1.18	1.18	1.06		1.74	
GER[1]			1.27	1.56		1.46	
MEL A						1.82	
MEL B	1.47	1.58	1.56	1.62	1.73	1.73	
NAML	2.09	2.03	2.15	2.02	1.97	1.99	

Significant changes between consecutive periods are indicated by asterisk, where ** indicates significant on five percent level and * on ten percent level (two-sided). The last column shows significant differences between the first and last period.
[1] Not tested for significant differences.

As a group, the Big Five leagues have, during the last 60 seasons, weakened their win dispersion little by little, from period to period. In overall, from the first to the last period, the difference is from 1.26 to 1.49 (weak significant). The ratio is not affected by including/excluding Germany, but the significance is stronger. Compared to the first period, which has the lowest ratio, it is the competitive balance in the periods from the middle of the 1970s to present that is significantly weakest. When including Germany,

[139] This is done in Chapter 6 when analysing single leagues (England and Norway).

the weakened win dispersion from period (III) to period (IV) is also significant. The insignificant (t-value = 1.14) weakened win dispersion between period (I) and (II) might indicate some negative effect on competitive balance, as a consequence of introducing UEFA team tournaments. The negative trend in the last thirty seasons might be related to commercialization, broadcasting and media, respectively, and creation and extension of the UEFA Champion League.

Much of the same trend can be found in the second tier of the Western European leagues, but at a weaker level of win dispersion (higher NSQF ratio). The Turkish and the Dutch leagues are not included in the first period.

On average, the 3WE leagues show a different trend, by moving a little bit up and down, and the last period is very similar to the first period. However, no significant difference between the periods can be found. This tier has a NSQF ratio that indicates better win dispersion than for the second tier, and compared with the Big Five at the end of the data period, these ratios are almost equal. However, early in the sample period, the 3WE leagues had weaker win dispersion than the Big Five leagues. The 2WE leagues are always weaker balanced than the 3WE leagues.

Among the Eastern European leagues that have existed during the whole sample period (EEW), the ratio has a U-form over the sample period. In the last period, among the Western European leagues, only the second tier has weaker NSQF ratio than these Eastern European leagues. Except from period (I), the last period is significantly weaker than the other periods, and especially the last three periods seem to have a negative trend. The changes between these periods are significant (period (IV) to (V) and (V) to (VI)).

The "old" Eastern European leagues have among the lowest NSQF ratios (best balanced) from period (II) to period (IV). Introduction of UEFA team tournaments might have had a different impact on win dispersion in the Eastern Europe leagues compared to the Big Five leagues. At least, all of these leagues improve win dispersion from the first to the second period.

One of the most interesting results in this analysis emerges from the restructuring of the leagues in Eastern Europe in the early 1990s. In the context of competitive balance, the political changes in this part of Europe give a unique opportunity to analyse the consequences of radical changes in the size of leagues. For win dispersion, the results are very strong. Average win dispersion is much weaker in the last period, when averaging (equally weighting) the "new" leagues, than in the periods where these leagues were parts of a bigger system. The effects on the NSQF ratio from the merger of the West and East German leagues are more difficult to interpret, except that the value in the last period is lower than in period (IV). If the difference between period (III) and (IV) is part of a trend towards weaker competitive balance, then the merge might be seen as positive in relation to win dispersion.

Also, the minor leagues operating during the whole sample period (MEL B) show weakened competitive balance in the win dispersion dimension, but only the difference between periods (I) and (V) is significant. The general level of competitive balance is slightly better than the second tier leagues, but when including more leagues in the last period (MEL A), the average is almost equal to the second tier leagues.

The general level of win dispersion in the Norwegian league, compared to the other leagues, is in the upper half during all periods. In addition, there is a U-form pattern in competitive balance over time in Norway, where the period around 1980 has the lowest value of the NSQF ratio. No significant differences between the periods are found.

In addition to Norway, a number of the other leagues will be shown below. First, competitive balance among the Big Five leagues is presented:

Table 4.2: Ten-season win dispersion measured by the NSQF ratio in the Big Five leagues from 1946/47 – 2005/06

	1946/47-1955/56 (I)	1956/57-1965/66 (II)	1966/67-1975/76 (III)	1976/77-1985/86 (IV)	1986/87-1995/96 (V)	1996/97-2005/06 (VI)	Sign I→VI
England	1.14	1.34**	1.44	1.46	1.44	1.61*	**
Spain	1.30	1.39	1.21*	1.33	1.47	1.38	
Italy	1.42	1.41	1.46	1.39	1.54*	1.67	**
France	1.20	1.24	1.22	1.45**	1.30*	1.30	
Germany			1.26	1.45	1.35	1.46	[1]

Significant changes between consecutive periods are indicated by asterisk, where ** indicates significant on five percent level and * on ten percent level (two-sided). The last column shows significant differences between the first and last period.

[1] No significant difference between period (III) and (VI).

During the last sixty seasons, the English top division had its strongest win dispersion period in the first ten season period. It is significantly better than all other periods. In addition, the last period (VI) has weakest win dispersion.[140] The French league is usually among the best balanced among the Big Five leagues, but the ten season period from the mid-70s to mid-80s, was an exception. This period is also significantly weaker than all the other periods. The Italian league is among the weakest balanced in all periods, and the last period is the weakest. It is difficult to find any trends of win dispersion in the Spanish league. The same can be said about the German league. These results are relatively similar to Bourg's (2004) estimates for parts (1980/81-1999/00) of the sample period used in this chapter, as well as for the surprise index in Groot (2008).

As shown earlier, the general level of win dispersion is weaker in the second tier of leagues in Western Europe. The table below shows the win dispersion for these leagues.

Table 4.3: Ten-season win dispersion measured by the NSQF ratio in the 2WE leagues from 1946(/47) – 2005/06

	1946/47-1955/56 (I)	1956/57-1965/66 (II)	1966/67-1975/76 (III)	1976/77-1985/86 (IV)	1986/87-1995/96 (V)	1996/97-2005/06 (VI)	Sign I→VI
Netherlands		1.41	1.99**	1.75**	1.73	1.89*	1
Scotland	1.56	1.77*	1.86	1.99	1.99	2.07	**
Portugal	1.70	1.89*	1.84	1.84	1.72	1.64	
Belgium	1.19	1.41	1.50	1.50	1.75	1.73	**
Turkey		1.49	1.20**	1.18	1.79**	1.75	1

Significant changes between consecutive periods are indicated by asterisk, where ** indicates significant on five percent level and * on ten percent level (two-sided). The last column shows significant differences between the first and last period.
[1] Significant increased win dispersion between period (II) and (VI) on five percent level.

By looking at each of the second tier European football leagues, interesting results can be found. For example, the Dutch league had a significant shift in competitive balance from period (II) to period (III). From the middle of the 1960s to the middle of the 1990s, the trend looks rather positive (reduced index). However, the last period (VI) had a significant reduction of competitive balance. These results are consistent with Koning (2000) and Groot (2008).

Among the other second tier leagues, the Scottish is among the weakest balanced in the first half of the sample, and has the highest index of these leagues in the second half. The trend looks negative, and different restructurings have not changed this trend.

[140] Possible explanations will be analysed in Chapter 6.

Another interesting observation is that the Turkish league had a significant strong negative shift in competitive balance from the middle of the 1980s.

As found in previous studies, competitive balance in NAML differs between the different sports. This is also shown in the table below:

Table 4.4: Ten-season win dispersion measured by the NSQF ratio in the NAML leagues from 1946(/47) – 2005(/06)

	1946/47-1955/56 (I)	1956/57-1965/66 (II)	1966/67-1975/76 (III)	1976/77-1985/86 (IV)	1986/87-1995/96 (V)	1996/97-2005/06 (VI)	Sign I→VI
AL	2.80	2.09**	1.83	1.73	1.59	2.05**	**
NL	2.23	2.09	1.81	1.73	1.68	1.80	**
MLB	2.46	2.04**	1.78*	1.81	1.62	1.90**	**
NBA	2.32	2.55	2.71	2.43	2.96**	2.77	**
NFL	1.60	1.56	1.70	1.51**	1.48	1.54	
NHL	1.99	1.96	2.42**	2.32	1.82**	1.74	

Significant changes between consecutive periods are indicated by asterisk, where ** indicates significant on five percent level and * on ten percent level (two-sided). The last column shows significant differences between the first and last period.

These results confirm that the National Football League (NFL) has the best win dispersion among the major leagues in all periods. However, the win dispersion in NFL is still weaker than for the average Big Five leagues, but the difference is smaller in the second half of the sample. Except from the period from the mid-60s to mid-70s, the ratio has been relatively stable. In the MLB, both leagues have much of the same trends, but in the periods with weaker win dispersion, the tendency has been that the American League (AL) is weaker balanced than the National League (NL). The NBA is the weakest, or second weakest, balanced in all periods, and especially in the last two periods (V and VI) the win dispersion in this league is much wider than for the other leagues. Win dispersion in the NHL has an inverse U-shape and seems to be in a positive trend in the second part of the sample.

The improvement in win dispersion in the first part of the sample for the MLB can be related to franchise moves, introduction of reverse-order-of-finish rookie draft in 1964, and the virtual disappearance of cash sales of star players (Quirk and Fort, 1992). According to Quirk and Fort these factors had a greater (absolute) impact on competitive balance than the negative effects of subsequent league expansions. Interestingly, in light of the 1994 strike in the MLB, both leagues seem to have weaker win dispersion in the post strike period (significant for the AL). This is consistent with the CBR measure of Humphreys (2002).

4.4.2. Championship Concentration

High concentration of championship winners is an indication of low turnover of top teams in a league, and represents one of the dimensions of competitive balance. The table below shows the empirical results of championship winners, measured by the HHI among the different groups of leagues.

Table 4.5: Ten-season championship concentration measured by the HHI in European football leagues and the NAML from 1946(/47) – 2005(/06)

	1946/47-1955/56 (I)	1956/57-1965/66 (II)	1966/67-1975/76 (III)	1976/77-1985/86 (IV)	1986/87-1995/96 (V)	1996/97-2005/06 (VI)	Sign I→VI
Big Five	0.20	0.27	0.34	0.33	0.30	0.39	**
2WE ex. Turkey	0.39	0.34	0.50	0.36	0.54**	0.42	
2WE	0.39	0.35	0.48	0.36	0.51	0.43	
3WE	0.36	0.28	0.30	0.28	0.32	0.44	
Norway[1, 4]	0.41	0.21	0.28	0.22	0.42	0.82	
EEW	0.38	0.43	0.36	0.26	0.32	0.34	
EEO	0.27	0.33	0.32	0.34			
SOV[4]	0.38	0.24	0.40	0.19		0.40	
CZE[4]	0.26	0.36	0.36	0.20		0.37	
YUG[2, 4]	0.26	0.34	0.28	0.28		0.37	
GER[4]	0.17	0.26	0.29	0.48		0.52	
MEL A						0.31	
MEL B	0.28	0.29	0.33	0.34	0.25	0.31*	
NAML[3]	0.32	0.44	0.23	0.22	0.22	0.21	*

Significant changes between consecutive periods are indicated by asterisk, where ** indicates significant on five percent level and * on ten percent level (two-sided). The last column shows significant differences between the first and last period.

[1] Adjusted for eight championships in period (I) and nine in period (II).

[2] Inclusive Bosnia-Herzegovina (1996/97-2005/06).

[3] MLB - World Series winners, NBA - Playoffs winners, NHL - Stanley Cup winners. NFL - NFC Championship Game winners between 1946 and 1965, and Super Bowl winners from 1966.

[4] Not tested for significant differences.

The general conclusions, on basis of the calculation of the HHI for the Big Five leagues during the sample period, are that the index was significantly better in the first period than the last, which might be reflecting a long run (negative) trend. Compared to other groups of leagues, including the NAML, the Big Five leagues had in general (much) lower index in the beginning of the sample period.

There is much variation in HHI from period to period, and therefore difficult to find trends. In Western Europe, the second tier has the highest HHI, except in the last period, which indicates that these leagues might have a higher degree of dominant teams. For leagues that have split, the same tendency as for win dispersion appears, where concentration is much higher on average in the last period than the previous average results.

In general, the championship concentration has been relatively stable in the NAML since the middle of the 1960s, and this level of the HHI is, on average, much lower than for the European leagues. The extension of post-seasonal championship playoffs may account for this difference. However, the Norwegian league had a (limited) post-seasonal playoff in the first period (I), and has a higher HHI than many of the other periods without playoff. These results might indicate that the introduction of the European Cup probably had very little significant impact on the distribution of finances between the teams in the Norwegian top level, which was based on pure amateurism before 1984. Another interesting result from the Norwegian league is the very high HHI in the last period (VI), as a consequence of the dominance by one team.

The dynasty period for Rosenborg BK might have been positively affected by the prize money becoming more significant. In 1995, Rosenborg BK won a fourth consecutive domestic championship and qualified for the UEFA Champions League. The sporting strength was now supplemented by a significant increase in the team's financial strength, and the team continued to win the domestic championship every season, until Vålerenga (Oslo) broke their run in the 2005 season. Apart from 2003, Rosenborg BK qualified for the UEFA Champions League all these seasons, including the 2005 season. Dynasties will tend to be reflected in higher HHI. Therefore it is appropriate to go further into some of the groups presented above (similar as for the win dispersion measure). Below, the HHI for the Big Five leagues is presented:

Table 4.6: Ten-season and 30-season championship concentration measured by HHI for the Big Five leagues from 1946/47 – 2005/06

	1946/47-1955/56 (I)	1956/57-1965/66 (II)	1966/67-1975/76 (III)	1976/77-1985/86 (IV)	1986/87-1995/96 (V)	1996/97-2005/06 (VI)	1946/47-1975/76 (VII)	1976/77-2005/06 (VIII)
England	0.16	0.16	0.16	0.52	0.20	0.38	0.10	0.20
Spain	0.26	0.54	0.42	0.26	0.42	0.30	0.34	0.27
Italy	0.22	0.28	0.22	0.40	0.32	0.42	0.20	0.28
France	0.22	0.22	0.54	0.20	0.30	0.32	0.18	0.14
Germany[1]	0.16	0.14	0.34	0.28	0.26	0.52	0.08	0.29

[1] Adjusted for nine championships in period (I) and 29 in period (VII).

After the Second World War, especially two periods have had dynasties in English football. Around 1980, Liverpool FC was the dominating team. After the formation of the English FA Premier League in 1992, Manchester United had a dynasty period. Both these periods (IV and VI) have much higher level of the HHI than the other periods. In general, the HHIs in the latter part of the data sample (VIII) are much higher than in the first half (VII). Also for the other leagues, HHI differs between the different periods. Only Italy has not had a period where HHI is higher than 0.5.

Real Madrid became the first winner of the European Cup in the 1955-56 season. This was a knock-out tournament for the league champions of the UEFA member associations. The format changed in 1992, when it became the UEFA Champions League (with Marseille as the first winner), but still only open to championship winners (www.uefa.com/competitions/ucl/history/index.html). Prior to the 1997/98 season two teams could qualify from the highest ranked leagues. This was increased by the addition of two teams from the 1999/00 season (based on uefa.com). The European Cup for championship winners and the UEFA Champions League formats might affect the domestic competition if the prize money is significant. This has been more significant after the introduction of the UEFA Champions League. The increased HHI between period (I) and (II) in the Big Five leagues, after the changes in the pan-European tournaments, is mainly driven by Spain.

The structural change in the (West) German league system had a negative influence on the HHI. The use of championship playoffs prior to the formation of the Bundesliga was associated with a greater dispersion of championship winners, similar to the experience in the NAML.

For the second tier leagues in Western Europe, the HHI for the different leagues is presented in the table below.

Table 4.7: Ten-season and 30-season championship concentration measured by HHI for the WE2 leagues and Norway from 1946(/47) – 2005(/06)

	1946/47-1955/56 (I)	1956/57-1965/66 (II)	1966/67-1975/76 (III)	1976/77-1985/86 (IV)	1986/87-1995/96 (V)	1996/97-2005/06 (VI)	1946/47-1975/76 (VII)	1976/77-2005/06 (VIII)
Netherlands	0.12	0.22	0.38	0.42	0.42	0.46	0.16	0.39
Scotland	0.36	0.32	0.68	0.36	0.82	0.50	0.29	0.40
Portugal	0.54	0.46	0.68	0.36	0.52	0.42	0.45	0.39
Belgium	0.54	0.36	0.30	0.30	0.42	0.30	0.34	0.31
Turkey[1]		0.38	0.34	0.38	0.36	0.46	0.34	0.26
Norway[2]	0.41	0.21	0.28	0.22	0.42	0.82	0.14	0.34

[1] Adjusted for eight championships in period (II) and 18 in period (VII).
[2] Adjusted for eight championships in period (I), nine in period (II) and 27 in period (VII).

Many of these leagues have the characteristic that a certain number of teams dominate the leagues season after season. For example in Portugal, during the whole sample period, three teams shared all but one championship (SL Benfica, 25, FC Porto 18, Sporting CP, 16), with Boavista FC in the 2000/01 season being the only other team to win the Portuguese league in the last 60 years. Also for the other leagues, particular teams have had a dominant share of the total championships, such as Celtic and Rangers in the Scottish league. This is also reflected in the calculations of HHI for the second tier of Western European leagues, where only two periods in the Dutch league (I and II) had a HHI lower than 0.30. Note that also HHI has a negative change between period (II) and (III) in this league, similar to the results from win dispersion.

As indicated above, it seems like the distribution of championships in general for the NAML had a positive shift on the HHI from period (II) to period (III). Therefore, it is appropriate to have a closer look at the different major leagues, as is presented in table 4.8.

Table 4.8: Ten-season and 30-season championship concentration measured by HHI for the NAML from 1946(/47) – 2005(/06)

	1946/47-1955/56 (I)	1956/57-1965/66 (II)	1966/67-1975/76 (III)	1976/77-1985/86 (IV)	1986/87-1995/96 (V)	1996/97-2005/06 (VI)	1946/47-1975/76 (VII)[2]	1976/77-2005/06 (VIII)[3]
MLB[1]	0.40	0.28	0.18	0.12	0.14	0.24	0.16	0.08
NFL	0.20	0.18	0.16	0.16	0.26	0.18	0.10	0.10
NBA	0.32	0.82	0.24	0.22	0.28	0.24	0.23	0.15
NHL[4]	0.36	0.46	0.34	0.36	0.18	0.19	0.29	0.13

[1] Adjusted for nine championships in period (V) and 29 in period (VIII).
[2] For the NFL, Championship winners NFL (1946/47-1975/76).
[3] For the NFL, Super Bowl winners (1976/77-2005/06).
[4] Adjusted for nine championships in period (VI) and 29 in period (VIII).

The NFL has a relatively low level of HHI in all periods. For the other leagues, more variation appears, but typically, the highest level of HHI is at the first two periods. The NBA in the second period (II) had a dominating team (Boston Celtics), winning nine out of ten finals. The team lost its domination, which is the main reason for the change in concentration mentioned above. NHL seems to have had a (positive) shift from period (IV) to period (V). In general, the two last periods (V and VI) for all NAML have relatively low HHI, indicating high spread of championship wining teams. The differences between the leagues are relatively small, especially for the last period, with only 0.06 as the highest difference (between MLB and NFL). For the MLB, the championship winner concentration is higher in the post-strike period than the previous period, and follows the results from win dispersion.

Comparing period (VII) and (VIII), the NAML have increased the turnover of champions between the two halves of the sample (MLB, NBA and NHL), or been stable (NFL). Except from the French league (which has a dynasty tendency at the time of writing up this thesis), other West European leagues shown above have a much lower turnover than the NAML, when looking at the last thirty seasons (period VIII).

4.4.3. Performance Persistence

In general, greater changes from season to season in performance for teams are seen as a sign of better competitive balance. The table below summarises performance persistence, measured by the SRCC, for the European football leagues, as well as for the NAML, after the Second World War.

Table 4.9: Ten-season performance persistence measured by SRCC for European football leagues and the NAML from 1946(/47) – 2005(/06)

	1946/47-1955/56 (I)	1956/57-1965/66 (II)	1966/67-1975/76 (III)	1976/77-1985/86 (IV)	1986/87-1995/96 (V)	1996/97-2005/06 (VI)	Sign I→VI
Big F. ex. Germ.	0.46	0.50	0.54	0.55	0.54	0.58	
Big Five	0.46	0.50	0.51	0.56	0.54	0.57	
2WE ex. NL/Tur	0.57	0.64	0.66	0.73	0.62**	0.59	
2WE	0.57	0.60	0.67	0.66	0.64	0.63	
3WE ex. Greece	0.53	0.49	0.44	0.49	0.43	0.52	
3WE	0.53	0.49	0.50	0.53	0.48	0.55	
Norway	0.46	0.40	0.28	0.41	0.34	0.48	
EEW	0.42	0.63**	0.59	0.48	0.54*	0.61	**
EEO	0.52	0.56	0.58	0.60			
SOV[2]	0.56	0.66	0.53	0.48		0.60	
CZE[2]	0.54	0.46	0.61	0.57		0.53	
YUG[2]	0.72	0.57	0.64	0.52		0.60	
GER[2]			0.47	0.73		0.53	
MEL A						0.60	
MEL B	0.46	0.51	0.50	0.59*	0.63	0.62	*
NAML[1]	0.57	0.60	0.62	0.59	0.48	0.50	

Significant changes between consecutive periods are indicated by asterisk, where ** indicates significant on five percent level and * on ten percent level (two-sided). The last column shows significant differences between the first and last period.

[1] SRCC in NAML are calculated on basis of rank, using on win percent in the given season. If two teams have the same win percent, the rank is further decided on basis on ranking in division. Washington Senators is split into two; numbers I and II. This means that from 1960 to 1961 Washington Senators (I) is relocated to Minnesota Twins, while Washington Senators (II) is treated as a new team.
Regarding Dallas Texans I and Baltimore Colts, "NFL considers the Texans and Colts to be separate teams" (en.wikipedia.org/wiki/Indianapolis_Colts). In addition, new teams are treated as lowest ranked in the previous season (for calculation purposes).

[2] Not tested for significant differences.

It is not easy to find general trends from the calculation of performance persistence in European football. Most of the differences between the periods are insignificant. Even if it seems like most of the groups have higher value (more persistence) on the Spearman's rank correlation coefficient (SRCC) in the last period (VI) compared to the first (I), only the EEW and the MEL B groups of leagues get this increase to be significant. The latter of these seems to be the only group of leagues with a kind of trend (towards higher persistence). The NAML might have a small positive shift between period (IV) and (V) (insignificant).

Also for performance persistence, the second tier leagues in Western Europe have weaker values than the other Western European leagues, as well as in comparison to most of the other groups of leagues. Relatively similar values between the Big Five and third tier are found. It does not seem that going from one bigger league to averaging smaller leagues in Eastern Europe has the same negative effect on performance persistence, as for the other dimensions of competitive balance.

As for the other measures, for some of the groups the different leagues are presented below.

Table 4.10: Ten-season performance persistence measured by SRCC for the Big Five leagues from 1946/47 – 2005/06

	1946/47- 1955/56 (I)	1956/57- 1965/66 (II)	1966/67- 1975/76 (III)	1976/77- 1985/86 (IV)	1986/87- 1995/96 (V)	1996/97- 2005/06 (VI)	Sign I→VI
England	0.35	0.36	0.51**	0.48	0.44	0.62**	**
Spain	0.62	0.63	0.56	0.67	0.55*	0.59	
Italy	0.45	0.57	0.70**	0.59*	0.66	0.64	
France	0.42	0.44	0.39	0.47	0.49	0.47	
Germany			0.41	0.62**	0.58	0.53	1

Significant changes between consecutive periods are indicated by asterisk, where ** indicates significant on five percent level and * on ten percent level (two-sided). The last column shows significant differences between the first and last period.
[1] No significant difference between period (III) and period (VI).

There are variations in SRCC over time for the different Big Five leagues. Especially the English, and maybe also the Italian, have indications of negative shifts in SRCC over the periods (insignificant between first and last period for Italy), while the other leagues seem to be more "stable" over time, without particular trends.[141] For England, the latter period has significantly higher performance persistence than any of the other periods. In general, the French league has lowest or second lowest persistence in all periods. These results look relatively consistent with Groot (2008).

Table 4.11: Ten-season performance persistence measured by SRCC for the WE2 leagues from 1946(/47) – 2005/06

	1946/47- 1955/56 (I)	1956/57- 1965/66 (II)	1966/67- 1975/76 (III)	1976/77- 1985/86 (IV)	1986/87- 1995/96 (V)	1996/97- 2005/06 (VI)	Sign I→VI
Netherlands		0.51	0.73*	0.66	0.68	0.76*	1
Scotland	0.63	0.64	0.68	0.76	0.69	0.56	
Portugal	0.61	0.71	0.71	0.78	0.64**	0.62	
Belgium	0.48	0.58	0.58	0.64	0.53	0.59	
Turkey		0.56	0.65	0.46**	0.67**	0.64	2

Significant changes between consecutive periods are indicated by asterisk, where ** indicates significant on five percent level and * on ten percent level (two-sided). The last column shows significant differences between the first and last period.
[1] Significant difference from period (II) to period (VI) on 1 percent level.
[2] No significant difference from period (II) to period (VI).

[141] Possible explanations will be analysed in Chapter 6.

For the second tier of leagues in Western Europe, it is difficult to find any clear trend during the sample period. Significant deviations between some of the periods seem to be more "randomly distributed", such as period (II) in the Dutch league. Some of these leagues have had periods with very high performance persistence, where SRCC has been above 0.70. Looking at the second part of the data sample, there might be indications that the Scottish and the Portuguese leagues are in a trend towards weakened persistency (improved competitive balance). The Belgian top division seems to be the one with lowest persistence in this tier of leagues. The performance persistence for the Dutch league is consistent with Groot (2008).

It might be interesting to compare Big Five leagues and second tier leagues in the last period (VI). England has higher performance persistence than Scotland. Apart from France and the Netherlands, the values are relatively equal between the leagues categorised in Big Five and second tier.

Performance persistence in the different NAML is presented in the table below:

Table 4.12: Ten-season performance persistence measured by SRCC for the NAML from 1946(/47) – 2005(/06)

	1946/47-1955/56 (I)	1956/57-1965/66 (II)	1966/67-1975/76 (III)	1976/77-1985/86 (IV)	1986/87-1995/96 (V)	1996/97-2005/06 (VI)	Sign I→VI
AL	0.70	0.73	0.35**	0.55	0.25**	0.61**	
NL	0.67	0.55	0.55	0.45	0.22*	0.42**	**
MLB	0.70	0.67	0.54*	0.54	0.26**	0.53**	**
NBA	0.43	0.66*	0.65	0.57	0.71**	0.62	
NFL	0.46	0.40	0.59**	0.51	0.41	0.29*	
NHL	0.68	0.67	0.72	0.75	0.55**	0.55	

Significant changes between consecutive periods are indicated by asterisk, where ** indicates significant on five percent level and * on ten percent level (two-sided).

Also for the NAML, it is difficult to find clear trends, even if there might be indications of a significant shift in the NHL between period (IV) and period (V) (but period (V) is not significantly different from period (I) and period (II)), and the NFL has in the last four periods been in a positive trend, from 0.59 in period (III) to 0.29 in period (VI). The baseball leagues are also, for this competitive balance dimension, weakened in the last decade. This might be surprising, since the strike in 1994 resulted in both luxury taxes and more revenue sharing, but it follows the results from both win dispersion and championship concentration.

4.4.4. General Comments

There is a significant division between the Big Five leagues and other leagues with regards to competitive balance. This is interesting when comparing European football leagues and NAML, because typically, the Big Five leagues are over-represented in samples analysing European football, and might therefore give a skewed picture of the general competitive balance in European football. Compared to the averaged within-seasonal NSQF ratio, it seems like the European Big Five football leagues have better win dispersion than the NAML, while the new leagues in the former Soviet Union have, on average, weaker win dispersion. These findings adjust the general opinion of the difference in competitive balance between the continents. Comparisons between NAML and Big Five on performance persistence show only small differences, from a situation where the Big Five is slightly better in the beginning of the period to a trend where NAML is slightly better than the Big Five at the end of the period. Prize concentration is better (more dispersed) in NAML.

4.4.5. Competitive Balance in European Football in the Last Decade

The following table shows the calculated values and rankings of competitive balance for each European football league during the last period of the data sample: 1996(97)-2005(/06). In general, based on rankings, the level of competitive balance is stated as A to E, where all have ten leagues, except from HHI that has 11 on rank C. This means that the top ten ranking is graded A, while the ranking from number 11 to 20 is graded B etc. This is done for each of the three competitive balance measures.

Table 4.13: Overview over competitive balance in three dimensions in European football leagues over the period 1996(/97) – 2005(/06)

League	NSQF		SRCC		HHI		League	NSQF		SRCC		HHI	
Albania	1.621	B	0.582	C	0.420	D	Kazakhstan	2.304	E	0.553	B	0.240	A
Andorra	2.218	E	0.634	D	0.220	A	Latvia	2.728	E	0.536	B	0.820	E
Armenia	2.506	E	0.632	D	0.320	C	Lithuania	2.413	E	0.780	E	0.420	D
Austria	1.656	C	0.615	D	0.200	A	Luxemburg	1.842	C	0.616	D	0.420	D
Azerbaijan	2.133	E	0.544	B	0.291	B	Macedonia	1.842	C	0.588	C	0.220	A
Belarus	2.078	D	0.652	E	0.150	A	Malta	2.084	E	0.829	E	0.300	B
Belgium	1.729	C	0.594	C	0.300	B	Moldova	2.341	E	0.728	E	0.460	D
Bosnia-Herz	N/A		N/A		0.180	A	The Netherlands	1.886	D	0.760	E	0.460	D
Bulgaria	2.060	D	0.772	E	0.300	B	Northern Ireland	1.723	C	0.509	A	0.280	B
Croatia	1.545	B	0.533	B	0.460	D	Norway	1.406	A	0.478	A	0.820	E
Cyprus	2.157	E	0.762	E	0.340	C	Poland	1.509	A	0.507	A	0.400	C
Czech rep	1.465	A	0.588	C	0.540	E	Portugal	1.640	C	0.618	D	0.420	D
Denmark	1.573	B	0.590	C	0.340	C	Romania	1.573	B	0.597	C	0.380	C
England	1.615	B	0.624	D	0.380	C	Russia	1.603	B	0.522	B	0.440	D
Estonia	2.548	E	0.751	E	0.322	C	San Marino	1.859	D	0.499	A	0.220	A
Faroe Islands	1.829	C	0.589	C	0.280	B	Scotland	2.071	D	0.560	B	0.500	E
Finland	1.457	A	0.410	A	0.280	B	Serbia Mont.	2.022	D	0.611	D	0.420	D
France	1.304	A	0.471	A	0.320	C	Slovakia	1.626	B	0.465	A	0.200	A
Georgia	1.987	D	0.400	A	0.360	C	Slovenia	1.565	B	0.649	E	0.580	E
Germany	1.456	A	0.533	B	0.520	E	Spain	1.381	A	0.594	C	0.300	C
Greece	2.023	D	0.730	E	0.820	E	Sweden	1.289	A	0.517	A	0.180	A
Hungary	1.608	B	0.575	B	0.200	A	Switzerland	1.434	A	0.395	A	0.280	B
Iceland	1.233	A	0.547	B	0.280	B	Turkey	1.748	C	0.635	D	0.460	E
Ireland	1.638	B	0.634	D	0.260	B	Ukraine	1.905	D	0.526	B	0.580	E
Israel	1.719	C	0.576	C	0.320	C	Wales	2.030	D	0.579	C	0.540	E
Italy	1.671	C	0.638	D	0.420	D							

There are big differences between the best and the weakest leagues, when it comes to competitive balance, for all three dimensions. The best balanced European football league, when taking into account all three dimensions for the last decade, is the Swedish. This is the only league to be ranked A in all three dimensions, while Slovakia, Switzerland and Finland have A in two dimensions and B in the third. At the bottom, Greece, Lithuania and Moldova have two E's and one D. When it comes to the Big Five leagues, the French (A, A, C) and the Spanish (A, C, C) are best balanced, while Germany is in the middle (A, B, E). Weakest are the English (B, D, C) and the Italian (C, D, D). Norway is one of the leagues having the same pattern as Germany, with one dominating team, and the two other dimensions are well balanced (A, A, E).

Three of the leagues have had relatively extreme dynasties during the sample period. These are the Norwegian (Rosenborg BK), the Latvian (Skonto Riga) and the Greek (Olympiakos). These leagues have all one team winning all championships, but one. An interesting observation can be made looking at the NSQF ratio for these three leagues. While both Greece and Latvia have a ratio above two, and Latvia has the weakest ratio of all European football leagues, the picture for the Norwegian league is the opposite, since only nine leagues have had better ratio on win dispersion. Also on performance persistence, Norway is at the A level. An interesting observation in this table is that the league with the weakest average NSQF ratio and HHI, Latvia, actually is much better balanced when the SRCC is used as measure. These differences illustrate the importance, as indicated earlier, of taking different dimensions into account when categorising a league's competitive balance. More thorough analyses of the single leagues are required to get a better overview over the reasons why the different measures give these results.

4.4.6. Multiple Match Outcomes

Many leagues divide from the simple league context by having multiple outcomes. Tie matches are a relatively frequent outcome in European football. From the examples in the table below, it seems like more than 20 % of the matches end as ties.

Table 4.14: The percent of matches ending as ties in a number of European football leagues

League	Period	Tied matches (%)
Big Five leagues		
England	1946/47-2005/06	25.9 %
Germany	1963/64-2005/06	26.0 %
Spain	1946/47-2005/06	24.4 %
Italy	1946/47-2005/06	32.1 %
France	1946/47-2005/06	26.4 %
Other selected European football leagues		
Turkey	1959-2005/06	29.1 %
Scotland	1946/47-2005/06	22.8 %
Portugal	1946/47-2005/06	24.0 %
Netherlands	1956/57-2005/06	25.8 %
Sweden	1946/47-2005	25.7 %
Norway	1948/49-2005	23.5 %

Among the European football leagues above, the Scottish has the lowest rate (22.8 %) of tied matches, while close to a third of the matches in the Italian top division have ended without a winner. Explanations of the differences in the share of tied matches between leagues are beyond the scope of this thesis.[142]

[142] This can be an interesting topic in a joint research with football science. Is it a result from differences in the traditional way of playing football (differences in playing style between leagues)? Finding indicators of the "traditional style of play", for example by measuring the number of goals scored,

Point score systems have also differed between leagues, as well as within leagues over time. The traditional (2,1,0) score system in European football was the general system, until the 1980's, when some leagues started to use alternative systems. England introduced the (3,1,0) system into its league in the 1981/82 season, and during the 1990s all leagues have converted to this system. By the 2005(/06) season, all UEFA member leagues operate the (3,1,0) point score system.[143] Therefore, how to treat the scoring system in the NSQF ratio is a significant topic for European football. Below, score systems from some of the leagues (during the sample period) are presented.

Table 4.15: Score systems for a number of European football leagues

League	(2,1,0) system	(3,1,0) system	Others
Big Five leagues			
England	Until 1980/81	1981/82 onwards	
Germany(W)	Until 1994/95	1995/96 onwards	
Spain	Until 1994/95	1995/96 onwards	
Italy	Until 1993/94	1994/95 onwards	
France	Until 1972/73, 1976/77-1987/88, 1989/90-1993/94	1988/89, 1994/95 onwards	1973/74-1975/76[a]
Other selected European football leagues			
Turkey	Until 1986/87	1987/88 onwards	
Scotland	Until 1993/94	1994/95 onwards	
Portugal	Until 1994/95	1995/96 onwards	
Netherlands	Until 1994/95	1995/96 onwards	
Sweden	Until 1989	1990 onwards	
Norway	Until 1986	1988 onwards	1987[b]

Sample period: After Second World War.
[a] Wins by three or more goals = 3, other wins = 2, draws= 1, losses = 0.
[b] Wins = 3, draw and shoot-out winner =2, draw and shoot-out loser =1, losses =0.

From Chapter 3, recall that the NSQF = ASD/ISD. According to Quirk and Fort (1992) and Cain and Haddock (2006), ISD(100%,0%) = $\frac{0.5}{\sqrt{m}} = \frac{1}{\sqrt{4m}}$. Assuming a three match outcome with uniform probabilities, the discussions in Chapter 3 showed that ISD(100%,50%,0%) = $\frac{1}{\sqrt{6m}}$. Given that ASD is based on the (100%,50%,0%) approach, the relative difference between the NSQF with and without tie-adjusted ISD will be:

incentives by score systems (offensive defensive play), and the general level of competitive balance (weaker competitive balance due to greater differences between the teams and less draw matches) might be relevant starting points of such analysis.

[143] Based on my observations on the data sample for this thesis.

$$NSQF[ISD(100\%,50\%,0\%)]_{uniform} / NSQF[ISD(100\%,0\%)]$$

$$= \frac{ASD}{1/\sqrt{6m}} \Big/ \frac{ASD}{1/\sqrt{4m}} = \sqrt{6/4} \approx 1.2247.$$

In other words, when using uniform probabilities in the ISD, the values are about 22.47 percent higher than the traditional calculations of the NSQF ratio in, for example, European football. If applying the (¼-½ -¼) probabilities, the relative difference is $\sqrt{2} \approx 1.414$.

These effects are shown in the table below, when comparing the "traditional" NSQF ratios between the English top division, based on different ways of measuring the idealized standard deviation, and the American League (AL) and the National League (NL) in the MLB.

Table 4.16: Comparison of NSQF ratios in the MLB and the top division in English football for alternative valuation of draw outcome

	MLB		English top division in football		
	AL	NL	Alt I	Alt II	Alt III
1976/77-85/86	1.743	1.728	1.465	1.794	2.072
1986/87-95/96	1.588	1.685	1.440	1.763	2.036
1996/97-05/06	2.053	1.799	1.615	1.978	2.284

Alt I = ISD(100%,0%); Alt II = $ISD(100\%,50\%,0\%)_{uniform}$;
Alt III = $ISD(100\%,50\%,0\%)_{(¼-½-¼)}$.

As can be seen in the table above, by applying ISD(100%,0%), the NSQF ratio for the English league has a better win dispersion than the NL and the AL (Alt I). However, when applying the tie-adjusted ISD, these differences are "wiped out", indicating a relatively similar level of win dispersion in the English top division and the MLB (for Alt II). For Alt III the result is opposite.

The "tie adjusted" NSQF ratio, calculated by Cain and Haddock for the top two divisions in English football, follows the pattern above, and is higher than the "traditional" ratios (their figure 2). In other words, when including ties as a match outcome in the idealized standard deviation, it is expected to improve (smaller dispersion), and hence, to increase the "adjusted" NSQF ratio (showing weaker win dispersion). However, Cain and Haddock do not comment on this important part of their results, and instead focus on explaining the time-series development in the different leagues. It would have been interesting to see Cain and Haddock's comments on a cross-sectional comparison, because their method of calculating the ("tie adjusted") NSQF ratio is an empirical innovation. The time-series changes in the NSQF ratio will also be reflected in the "traditional" NSQF ratio. The results from the tie-adjusted NSQF ratios above and Cain and Haddock (2006) indicate that tied matches in European football might be one of the explanations of better win dispersion for (some of) these leagues. By changing the way of calculating the idealized standard deviation, the general view that European football leagues are better balanced (i.e. better win dispersion) than the NAML (except the NFL) might need to be reconsidered.

4.4.7. Playoff Concentration

The simple league measure for prize concentration earlier in this chapter was only related to championship winners. In this section, prize concentration will be expanded to concentrate on championship playoffs in NAML and Australian rugby, as well as for qualifications for "post-seasonal" UEFA Champions League among the Big Five leagues in European football, and the Scandinavian Royal League. The last two are chosen instead of leagues with championship playoffs in Europe because, as will be shown in Chapter 5, including championship playoff appears to be just a temporary "experiment" for many leagues, with only the San Marino league having had a continuous system over the analysed period. Therefore, it seems to be difficult to find a period that is suitable for comparative analyses with sports on other continents. But increases in number of teams from some leagues qualifying for the UEFA Champions League, although a relatively new phenomenon, provide possibilities for analysis for the last ten seasons.

Both NAML and Australian leagues have long traditions of post-seasonal championship playoffs. This gives opportunities to apply prize concentration measures related to complex league systems. The analyses below will be done by applying HHI for championship playoff concentration in the four NAML and two leagues in Australia. Applying the measure on the NAML, at least three issues have to be taken into account. First, the number of teams qualifying for the playoffs differs, both between leagues and within leagues, over time. Second, the number of teams in the different leagues has increased during time. Third, league tournament structure has changed significantly over time in all leagues. Therefore, the "pure" HHI has to be adjusted, as shown in Chapter 3, into the GK PO-ratio. Also the PO-deviation measure will be used in these analyses.

4.4.7.1. North American Major Leagues

The GK PO-ratios for the playoffs in NAML are shown in the two tables below.

Table 4.17: Ten-season distribution of playoff qualifiers in the NAML measured by the GK PO-ratio

League	1966/67-1975/76	1976/77-1985/86	1986/87-1995/96	1996/97-2005/06
NFL[1]	1.830	1.380	1.295	1.282
MLB[2]	2.641	2.241	1.750	2.003
NBA	1.326	1.248	1.276	1.203
NHL[3]	1.300	1.160	1.143	1.307

[1] NFL is included AFL prior to the merger; [2] No playoff in 1994; [3] No competition in 2004/05.

Table 4.18: Five-season distribution of playoff qualifiers in the NAML measured by the GK PO-ratio

	66/67-70-71	71/72-75/76	76/77-80-81	81/82-85/86	86/87-90-91	91/92-95/96	96/97-00-01	01/02-05/06
NFL[1]	2.192	2.067	1.925	1.426	1.630	1.527	1.444	1.474
MLB[2]	2.835	3.840	3.865	1.885	2.326	2.763	2.173	2.340
NBA	1.420	1.396	1.350	1.274	1.365	1.341	1.393	1.323
NHL[3]	1.220	1.364	1.286	1.184	1.096	1.267	1.459	1.377

[1] NFL is included AFL prior to the merger; [2] No playoff in 1994; [3] No competition in 2004/05.

MLB has the highest playoff concentration in all periods. When using the ten season period, the post strike period is more concentrated than the previous period. This is consistent with Hadley et al. (2005). The NFL seems to have had a shift towards more dispersion of qualifying teams from the early 1980s, compared to the previous periods. The NHL is best balanced in the first three periods, while both NFL and NBA are slightly better in the last period. The NBA and NHL are relatively similar during the whole sample period, while the NFL is also at the same level in the two last decades. Actually, the NBA is the best balanced in the last part of the sample. This is surprising, related to general competitive balance calculations earlier in this chapter, and might make one reconsider the relatively unambiguous view on competitive balance in the NBA. Instead, the prize concentration might rather be a problem in the MLB. This is interesting in relation to the Blue Ribbon panel on Baseball Economics for Major League Baseball, which focused on playoff qualification as the principal indicator of competitive balance. This is another indication that reflecting competitive balance through a single measure is difficult. Not only must the three simple league dimensions be taken into account, but additional information related to complex league structure might also be relevant pieces in the total picture of competitive balance.

Related to the discussion in Chapter 3, one obvious determinant to analyse when it comes to drivers for playoff concentration, is whether leagues with relatively few playoff places are treated appropriately by this measure. If the share is low, it means that dominating teams bias the distribution relatively more compared to the idealized distribution, because in the GK PO-ratio the actual HHI is calculated only on teams qualifying, while the idealized takes into account all teams in the league. The two tables below show the number of playoff places as a proportion of the total number of teams in a league for each major league over each ten-year and five-year period, respectively.

Table 4.19: Ten-season relative number of playoff places compared to the total number of teams in the NAML

	1966/67-1975/76	1976/77-1985/86	1986/87-1995/96	1996/97-2005/06
NFL[1]	0.284	0.364	0.397	0.386
MLB[2]	0.149	0.171	0.168	0.270
NBA	0.545	0.584	0.611	0.548
NHL[3]	0.609	0.740	0.699	0.560

[1] NFL is included AFL prior to the merger; [2] No playoff in 1994; [3] No competition in 2004/05.

Table 4.20: Five-season relative number of playoff places compared to the total number of teams in the NAML

	66/67-70-71	71/72-75/76	76/77-80-81	81/82-85/86	86/87-90-91	91/92-95/96	96/97-00-01	01/02-05/06
NFL[1]	0.260	0.308	0.329	0.400	0.371	0.423	0.395	0.377
MLB[2]	0.130	0.167	0.156	0.185	0.154	0.185	0.274	0.267
NBA	0.597	0.506	0.541	0.626	0.640	0.584	0.552	0.544
NHL[3]	0.643	0.585	0.716	0.762	0.762	0.645	0.584	0.533

[1] NFL is included AFL prior to the merger; [2] No playoff in 1994; [3] No competition in 2004/05.

The MLB has a much lower share of teams to playoffs, while both NBA and NHL have more than half of the teams to playoffs. This means that domination in MLB will be more highly weighted than domination in the NHL, because of the use of squared differences. The correlation, both in total and within each league, in relation to the share of teams to playoff and the GK PO-ratio, might give further information on these relationships, and is presented in the table below.

Table 4.21: Five- and ten-season correlation coefficient between the GK PO ratio and the relative number of playoff places compared to the total number of teams in the NAML

	Five-season correlation	Ten-season correlation
NFL[1]	-0.949	-0.990
MLB[2]	-0.465	-0.400
NBA	-0.318	0.020
NHL[3]	-0.844	-0.940
NAML	-0.834	-0.842

[1] NFL is included AFL prior to the merger; [2] No playoff in 1994; [3] No competition in 2004/05.

There is a strong correlation between the number of qualifying teams and the index in the NAML, which confirms the expected relationship. In particular, the NFL and the NHL have a strong correlation between the index and the numbers of teams that qualify for playoffs. On the other hand, the correlation is smaller in the two leagues with results that might look unusual, the MLB and, especially, the NBA (where the latter has no correlation). The table below presents the calculations based on the alternative measure (PO deviation) from Chapter 3.

Table 4.22: Ten-season distribution of playoff qualifiers in the NAML measured by PO deviation

	1966/67-1975/76	1976/77-1985/86	1986/87-1995/96	1996/97-2005/06
NFL[1]	13.159	11.956	11.091	10.866
MLB[2]	7.331	8.829	5.571	11.908
NBA	9.577	13.548	14.779	12.936
NHL[3]	9.367	11.240	9.039	13.396

[1] NFL is included AFL prior to the merger; [2] No playoff in 1994; [3] No competition in 2004/05.

This measure can be adjusted by the number of participants (teams) in the league during the specific period, where one team might have ten observations during a period of ten seasons (adj. PO deviation). The results are presented in table 4.23.

Table 4.23: Ten-season distribution of playoff qualifiers in the NAML measured by the adjusted PO deviation over the period 1966(/67) – 2005(/06)

	1966/67-1975/76	1976/77-1985/86	1986/87-1995/96	1996/97-2005/06
NFL[1]	0.051	0.043	0.039	0.035
MLB[2]	0.032	0.034	0.023	0.040
NBA	0.062	0.060	0.056	0.044
NHL[3]	0.068	0.056	0.039	0.052

[1] NFL is included AFL prior to the merger; [2] No playoff in 1994; [3] No competition in 2004/05.

Table 4.24: Five-season distribution of playoff qualifiers in the NAML measured by the adjusted PO deviation over the period 1966(/67) – 2005(/06)

	66/67-70-71	71/72-75/76	76/77-80-81	81/82-85/86	86/87-90-91	91/92-95/96	96/97-00-01	01/02-05/06
NFL[1]	0.056	0.063	0.060	0.050	0.055	0.057	0.047	0.045
MLB[2]	0.036	0.057	0.052	0.034	0.035	0.037	0.054	0.057
NBA	0.086	0.071	0.067	0.067	0.067	0.071	0.064	0.055
NHL[3]	0.072	0.077	0.071	0.071	0.052	0.059	0.070	0.060

[1] NFL is included AFL prior to the merger; [2] No playoff in 1994; [3] No competition in 2004/05.

These results are different from the previous analyses. The MLB is now relatively better balanced, and has a better relative distribution in the first three periods, while the NFL has the best index in the last period. In general, the NBA and the NHL are worst balanced. As indicated, these results might be driven by the number of teams to qualify for playoffs.

4.4.7.2. Australian Leagues

As an alternative, two leagues from Australia are chosen for the championship playoff (semi-finals) analyses. These leagues are the Australian National Rugby League (ARL/NRL)[144] and the Australian Rules Football League (VFL/AFL).[145] They are, in general, organised as closed one division leagues.[146] The top k teams qualify for championship playoffs, called semi-finals, where the match schedule is based on a seeding system related to the regular season performance. On policy, these leagues might be closer to the NAML, as they for example are closed. The following tables

[144] More information about this league can be found at www.nrl.com.
[145] More information about this league can be found at www.afl.com.au.
[146] Except from the 1997 season, when the ARL had a rival competition from the Super League. More information about the "Super League war" can be found in for example en.wikipedia.org/wiki/Australian_Rugby_League and related links.

show the number of teams qualifying for playoffs and the total number of teams in the leagues over different time periods.

Table 4.25: Ten-season descriptive data on playoffs in Australian leagues over the period 1966/67 – 2005/06

	1966/67-1975/76	1976/77-1985/86	1986/87-1995/96	1996/97-2005/06
VFL/AFL - teams to PO	44	50	60	80
VFL/AFL - teams total	120	120	144	160
VFL/AFL - PO share	0.367	0.417	0.417	0.500
ARL/NRL - teams to PO	43	50	53	86
ARL/NRL - teams total	118	126	158	167
ARL/NRL - PO share	0.364	0.397	0.335	0.515

Table 4.26: Five-season descriptive data on playoffs in Australian leagues over the period 1966/67 – 2005/06

	66/67-70-71	71/72-75/76	76/77-80-81	81/82-85/86	86/87-90-91	91/92-95/96	96/97-00-01	01/02-05/06
VFL/AFL - #PO	20	24	25	25	25	35	40	40
VFL/AFL - #tot	60	60	60	60	68	76	80	80
VFL/AFL - PO sh	0.333	0.400	0.417	0.417	0.368	0.461	0.5	0.5
ARL/NRL - #PO	20	23	25	25	25	28	46	40
ARL/NRL - #tot	58	60	60	66	74	84	93	74
ARL/NRL - PO sh	0.345	0.383	0.417	0.379	0.338	0.333	0.495	0.541

Both leagues have a pattern where the number of teams participating in the leagues has increased during time, as well as the number of teams reaching playoffs (semi-finals). Also the share of teams to playoffs has increased over time, and is at the end of the sample period at 50 percent in the VFL/AFL and 54.1 percent in the ARL/NRL. In general, the differences between the two leagues are much smaller than in the analyses of the NAML. The calculations from the GK PO-ratio are presented in the tables below.

Table 4.27: Ten-season distribution of playoff qualifiers in Australian leagues measured by the GK PO-ratio over the period 1966/67 – 2005/06

	1966/67-1975/76	1976/77-1985/86	1986/87-1995/96	1996/97-2005/06
VFL/AFL	1.624	1.570	1.336	1.191
ARL/NRL	1.782	1.553	1.549	1.474

Table 4.28: Five-season distribution of playoff qualifiers in Australian leagues measured by the GK PO-ratio over the period 1966/67 – 2005/06

	66/67-70-71	71/72-75/76	76/77-80-81	81/82-85/86	86/87-90-91	91/92-95/96	96/97-00-01	01/02-05/06
VFL/AFL	1.980	1.792	1.555	1.795	1.758	1.369	1.297	1.340
ARL/NRL	2.083	1.701	1.555	1.608	2.000	1.946	1.729	1.257

The Australian leagues have much of the same pattern as the NAML in playoff concentration. This is both related to time and level. Except from the 1990s, when the VFL/AFL seems to be much better balanced, the two leagues are relatively equal in the concentration ratios.

As for the NAML, the alternative playoff measure is also applied for the Australian leagues, and is presented below.

Table 4.29: Ten-season distribution of playoff qualifiers in the NAML measured by the adjusted PO deviation over the period 1966/67 – 2005/06

	1966/67-1975/76	1976/77-1985/86	1986/87-1995/96	1996/97-2005/06
VFL/AFL	0.084	0.084	0.061	0.058
ARL/NRL	0.091	0.074	0.057	0.062

Table 4.30: Five-season distribution of playoff qualifiers in the NAML measured by the adjusted PO deviation over the period 1966/67 – 2005/06

	66/67-70-71	71/72-65/66	76/77-80-81	81/82-85/86	86/87-90-91	91/92-95/96	96/97-00-01	01/02-05/06
VFL/AFL	0.095	0.103	0.090	0.102	0.085	0.069	0.075	0.073
ARL/NRL	0.101	0.093	0.090	0.076	0.083	0.073	0.077	0.059

The pattern between the Australian and the NAML are relatively similar when it comes to trends, but for this measure, the general level is higher in the Australian leagues. Taking into account the two measures of playoff concentration, there might be indications that the Australian leagues have a slightly higher concentration than the NAML.

4.4.7.3. UEFA Champions League Concentration

As mentioned earlier, it is much more difficult to analyse the prize to qualify for post-seasonal play in European football than in the NAML. However, because the UEFA Champions League has increased the number of teams that can qualify for the tournament beyond the domestic league champions, there is now an additional important sporting prize in the domestic leagues to be analysed. This mainly affects the bigger leagues, since the most of the smaller leagues, such as the Norwegian, are still limited to one qualifying team. The following analyses about UEFA Champions League concentration will therefore only concentrate on the biggest leagues in European football (Big Five), in a ten season period. The table below presents descriptive statistics related to UEFA Champions League in these leagues.

Table 4.31: Five- and ten-season distribution of possible UEFA Champions League qualifiers among the Big Five leagues over the period 1996/67 – 2005/06

	England				Germany				Italy				Spain				France			
		96/7-00/1	01/2-05/6	96/7-05/6		96/7-00/1	01/2-05/6	96/7-05/6		96/7-00/1	01/2-05/6	96/7-05/6		96/7-00/1	01/2-05/6	96/7-05/6		96/7-00/1	01/2-05/6	96/7-05/6
1	Man Utd	5	5	10	B.München	5	5	10	Juventus[1]	4	5	9	Real Madr.	5	5	10	Lyon	3	5	8
2	Newcastle	1	2	3	Bayer Lev	4	2	6	Parma	3	0	3	Barcelona	5	4	9	Monaco	2	3	5
3	Arsenal	4	5	9	Bor Dortmund	3	2	5	Inter Milan	2	5	7	Ath Bilbao	1	0	1	Paris SG	2	1	3
4	Chelsea	1	4	5	Kaiserslautern	1	0	1	AC Milan	2	5	7	Mallorca	2	0	2	Lille	1	2	3
5	Leeds Utd	1	0	1	Hertha Berlin	1	0	1	Lazio	3	1	4	Valencia	2	3	5	Lens	1	1	2
6	Liverpool	1	3	4	Hamburger	1	1	2	Fiorentina	1	1	2	Deportivo	2	3	5	Bordeaux	1	1	2
7	Everton	0	1	1	1860 München	1	0	1	Roma	1	2	3	Real Soc	0	1	1	Marseille	1	1	2
8					Schalke	1	1	2	Udinese C	0	1	1	RC Celta	0	1	1	Metz	1	0	1
9					VfB Stuttgart	0	1	1					Villarreal	0	1	1	Nantes	1	0	1
10					SV Werder B	0	3	3					Real Betis	0	1	1	Auxerre	0	1	1
11													Osasuna	0	1	1				
	Total to CL	13	20	33	Total to CL	17	15	32	Total to CL	16	20	36	Total to CL	17	20	37	Total to CL	13	15	28
	League tot	100	100	200	League tot	90	90	180	League tot	90	92	182	League tot	102	100	202	League tot	92	98	190
	% CL	13	20	16.5	% CL	18.9	16.7	17.8	% CL	17.8	21.7	19.8	% CL	16.7	20	18.3	% CL	14.1	15.3	14.7

[1] Juventus later lost its place in the UEFA Champions League after the domestic season 2005/06.

As can be shown from the overview over teams qualifying for the UEFA Champions League for the different leagues, the number of teams differs both between the two five season periods and between leagues. This is because of general expansions of teams to the tournament during the period, and that UEFA league rankings decide the number of teams to this post-seasonal play. Based on these data, the calculated GK PO-ratios are presented in the following table.

Table 4.32: Five- and ten season distribution of possible UEFA Champions League qualifiers in Big Five leagues measured by the GK PO-ratio over the period 1996/97 - 2005/06

	1996/97-2000/01	2001/02-2005/06	1996/97-2005/06
England	6.131	4.673	5.593
Italy[1]	3.808	5.132	4.344
Germany	4.150	4.241	4.017
Spain	5.017	3.756	4.408
France	2.977	4.498	3.714

[1] It is not taken into account that after the 2005/06 season a scandal resulted in relegation of the winner, and also that the runners-up got penalties (however, the runners-up was still given a qualification place for the UEFA Champions League).

Based on the ten season period, the English league is relatively more concentrated than the others, and the French is the least concentrated. However, when looking at the two five season periods, they have opposite trends. Note that the GK-PO concentration ratios for European football are much higher than for both the NAML and the Australian leagues.

The alternative measure (PO-deviation) for European football, represented by the Big Five leagues, adjusted by the number of observations during the analysis periods, is presented below.

Table 4.33: Five- and ten-season distribution of possible UEFA Champions League qualifiers in Big Five leagues measured by the adjusted PO deviation over the period 1996/97 – 2005/06

	1996/97-2000/01	2001/02-2005/06	1996/97-2005/06
England	0.354	0.571	0.833
Italy[a]	0.301	0.598	0.702
Germany	0.409	0.340	0.638
Spain	0.395	0.414	0.737
France	0.143	0.302	0.395

For the ten season period, the same relationship between the leagues can be found by the alternative measure. For the smaller periods, the results can differ from the other measure (as for the English FA Premier League). Similar to the GK PO-ratio, the alternative concentration ratio suggests much more concentration at the top of the European football leagues than for the leagues in the other continents analysed here.

As for the NAML, the correlations between the GK PO-ratio and UEFA CL-share are calculated in table 4.34.

Table 4.34: Correlation coefficients between the GK PO-ratio and the share of teams to possibly qualify for the UEFA Champions League in the Big Five leagues over the period 1996/97 – 2005/06

	Five-season correlation	Ten-season correlation
Big Five	-0.1038	0.0947

The conclusion from this simple correlation analysis is that the GK PO-ratio in European football is *not* correlated with the share of teams that qualify for the tournament.

In their analysis of the domination by the top teams in European football, Hoehn and Szymanski (1999) use a sample period of 1991-1997 for England. During this time 31 different teams competed in the top division; the number of European competition places in the period was 32, and 13 different teams appeared in European competitions (including the UEFA Cup and the UEFA Cup Winners' Cup). The top seven teams accounted for 81 percent of these European appearances, with 48 percent by the top three teams.

Regarding newly promoted teams and qualifications for the UEFA Champions League, for example the English FA Premier League had none of these teams qualified for this tournament during the ten season period.

4.4.7.4. Scandinavian Royal League

Because of the foundation of the Scandinavian Royal League (RL),[147] there is another opportunity to measure prize concentration for the domestic Scandinavian football leagues. The top four teams, based on the end-of-season table, play in this international post-seasonal tournament. Since the qualification structure is fixed, this seems, intuitively, to be a better "yardstick" to measure prize concentration than a combination including both qualifications for the UEFA Cup and the UEFA Champions League for these leagues. Descriptive data from the three domestic leagues, for the period where the Scandinavian Royal League has existed, is presented below.[148]

Table 4.35: Descriptive data and analysis of the concentration of teams qualifying for the Scandinavian Royal League over the period 20(03/)04 – 20(06/)07

	Norway		Sweden		Denmark	
	Teams	2004-2007	Teams	2004-2007	Teams	2003/4-2006/7
1	Vålerenga IF	3	Djurgårdens IF	3	FC København	4
2	SK Brann	3	IFK Göteborg	3	Brøndby IF	3
3	Lillestrøm	3	Halmstads BK	2	OB	3
4	Rosenborg BK	2	IF Elfsborg	2	FC Midtjylland	2
5	Viking	1	Kalmar FF	2	AaB	2
6	Stabæk	1	AIK Solna	1	Viborg FF	1
7	Tromsø IL	1	Hammarby IF FF	1	Esbjerg fB	1
8	Start	1	Helsingborgs IF	1		
9	Lyn Oslo	1	Malmö FF	1		
GK		2.273		2.288		2.444
PODev		0.293		0.230		0.403

These results suggest that the Danish league has highest playoff concentration. Evidence is related both to FC Copenhagen, as the only participant in all seasons, and a smaller number of teams being qualified for the Royal League in total. The post-seasonal Royal League qualification concentration ratio is based on four seasons, and can therefore only be treated as an indication of concentration.

As mentioned in Chapter 3, the difference between failure prize and success prize, the success prize - failure prize ratio, might be an indicator of the competitiveness of a league. Applying qualifications for the Scandinavian Royal League as the success

---

147 See www.royalleague.com .

148 The season 200(6/)7 is included, even though the tournament was cancelled. However, the cancellation was a reality after the domestic leagues were finished.

criteria and relegation (including the relegation playoff) as the failure prize threshold in the domestic leagues, the results are given in the tables below.

Table 4.36: Marginal points (pts) to qualify for the Royal League (RL) and to relegate (Rel) over the period 20(03/)04 – 20(06/)07

	Norway		Sweden		Denmark	
	Marg. pts to RL	Marg. pts to Rel	Marg. pts to RL	Marg. pts to Rel	Marg. pts to RL	Marg. pts to Rel
2004	40	27	41	28	57	27
2005	42	30	43	30	53	25
2006	44	29	42	22	54	26
2007	44	27	40	23	58	25

Table 4.37: Marginal win percent (WP) to qualify for the Royal League and to relegate, and the success prize – failure prize ratio (S-F ratio) over the period 20(03/)04 – 20(06/)07

	Norway			Sweden			Denmark		
	Marg. WP to RL	Marg. WP to Rel	S-F ratio[1]	Marg. WP to RL	Marg. WP to Rel	S-F ratio[1]	Marg. WP to RL	Marg. WP to Rel	S-F ratio[1]
2004	0.538	0.385	0.153	0.577	0.423	0.154	0.621	0.288	0.333
2005	0.577	0.423	0.154	0.596	0.423	0.173	0.576	0.288	0.288
2006	0.615	0.423	0.192	0.596	0.346	0.250	0.591	0.303	0.288
2007	0.615	0.365	0.250	0.577	0.346	0.231	0.621	0.288	0.333

[1] Success-failure ratio based on difference in win percentage.

Also by this measure, Denmark is weakest balanced, and confirms the descriptive statistics presented in the previous table. The two other leagues have a negative trend, where the two last seasons have greater difference than the two first periods.

4.4.8. Relegation-Promotion

One measure that can be applied in open leagues is the suggested rate of promoted teams to be relegated within two seasons, as presented in Chapter 3. The table below shows the indexes for the Big Five leagues and Norway from European football, for a period of 40 seasons.

Table 4.38: Two-season relegation ratio for promoted teams in Big Five leagues and Norway over the period 1966(/67) – 2005(/06)

	1966(/67) – 1975(/76)	1976(/77) – 1985(/86)	1986(/87) – 1995(/96)	1996(/67) – 2005(/06)
England	1.301	1.049	1.563	2.042
France	1.108	1.866	1.361	1.622
Germany	2.169	2.181	1.816	1.636
Italy	2.298	1.962	1.306	1.534
Spain	1.722	1.964	1.521	1.912
Norway	1.167	1.573	1.649	1.542

There is much variation from period to period, but a number of trends can be found. First, the top division in England has, from the mid-1970s, gone from having the best ratio, almost equal to the idealized probabilities, to a situation where promoted teams have more than twice the likelihood of relegation, compared to the expected shares in an idealized league. This is consistent with the general negative trend of other measures of competitive balance in the English top division. Interesting in relation to the increased gap between the two highest level divisions in English football from the middle of the 1980s, is that it coincides with the bigger differences between these divisions when it comes to broadcasting revenues (see also Chapter 6).

Italy and Germany have better ratios in the latter half of the sample than in the first half, when these leagues are worst balanced. The most recent period shows that France, Germany, Italy and Norway have similar ratios, while England and Spain have larger ratios indicating a bigger competitive gap in these two leagues between the top two divisions.

Summarising the empirical results in this chapter, NAML has in general weaker win dispersion than European football leagues. However, win dispersion has increased in the biggest Western European leagues, where six out of ten are significant. Only the NBA follows this pattern in the NAML. Comparing different groups of leagues, there is a tendency for bigger leagues to have a better win dispersion in European football, especially identified by comparing old Eastern European leagues with the average of the new smaller leagues. Moreover, comparing the smaller Western European leagues with the bigger gives the same pattern. However, the relationship between league size and win dispersion is not unambiguous, because of the relatively weak NSQF indexes in the second tier leagues (2WE).

Even if the Big Five leagues have significantly increased their championship concentration, opposite of the pattern in the NAML, the results are in general less clear for this and the performance persistence dimension of competitive balance. In general, championship concentration and performance persistence are calculated to be better in the NAML than European football leagues in the last decade of the sample. Also for playoff concentration, the results are better for the NAML than the Big Five leagues. In general, the upper part of the Big Five leagues is more concentrated than both the Scandinavian football leagues and the two selected Aussie leagues. The results also show that in some European leagues, competitive balance varies between the different dimensions, such as for the Norwegian league in the last decade, and that most measures indicate worsened competitive balance in the English top division over the sample period.

5. The Determinants of Competitive Balance across Leagues

5.1. Introduction

Chapters 2 and 3 went through conceptual issues on competitive balance and how it might be measured, both in a simple and in a complex league context. Chapter 4 gave an empirical overview over the level of competitive balance within and across both single leagues' and groups of leagues, as well as including empirical analyses by measures directly related to complex league structure. This chapter will focus on determinants of differences in competitive balance, especially across European football leagues.

Analyses of possible factors affecting competitive balance between the member leagues of the UEFA have a different challenge compared to similar analyses between the NAML. While the differences between the NAML to a large extent have been explained by differences in the level of market regulations, such analyses have rarely been done in Europe. One explanation is that the leagues among the UEFA members are relatively similarly organised, with few restrictions on "free market" based open leagues. The challenge is thus to find a different perspective to analyse differences in competitive balance between the leagues in the UEFA area.

The number of teams in the top division, rules for relegation, the number of teams qualifying for play-off etc., are issues that governing bodies have to take into account when deciding the tournament and prize structure. Effects of structural changes and differences in competition, in the context of competitive balance, will also be included in a cross-sectional regression analysis of determinants for differences in competitive balance across all European football leagues. The last part of the chapter will discuss championship playoff in the context of competitive balance in European football.

5.2. Literature Review

Even if many empirical studies on competitive balance has been related to single leagues (Berri et al., 2005), there are also a number of studies comparing competitive balance across leagues. The table below gives an overview over some of these studies.

The first contributions used different NAML as basis for this research. From the end of the 1990s comparisons among European leagues (usually football) have become more usual as well. Comparing competitive balance in European Football leagues with competitive balance in NAML is also part of this literature.

Table 5.1: Overview over literature concerning differences in competitive balance across leagues

Authors	***Team sports***
Quirk and Fort (1992)	NAML: NBA, MLB, NFL, NHL
Fort and Quirk (1995)	NAML: MLB, NBA, NFL, NHL
Vrooman (1995)	NAML: MLB, NBA, NFL
Hoehn and Szymanski (1999)	Football: Italy, Germany, Spain, England
Szymanski and Kuypers (1999)	Football: England, Scotland, Spain, Italy, the Netherlands
Dobson and Goddard (2001)	Football: England, Scotland, Spain, Italy, France, Germany
Szymanski and Smith (2002)	Football: Italy, Spain, Portugal, England NAML: NFL, MLB, NHL
Buzzacchi et al. (2003)	Football: England, Italy, Belgium NAML: MLB, NFL, NHL
Bourg (2004)	Football: France, Germany, England, Spain, Italy
Lee (2004)	Baseball: MLB, JPBL (CL, PL), KPBL
Gerrard (2004b)	Football: 16 Western European leagues
Dejonghe (2004)	Football: Spain, Italy, England, Germany, France, the Netherlands, Portugal, Scotland, Belgium
Berri et al. (2005)	Football: MLS, North American Soccer League, France, Spain, Germany, Italy, England American football: Canadian Football League, Arena Football League, NFL, AFL. Hockey: NHL, World Hockey Association Baseball: MLB Basketball: NBA, American Basketball Association
Szymanski and Valletti (2005b)	Football: England NAML: MLB
Gerrard (2006a)	Football: England, France, Germany, Italy, Spain
Groot (2008)	Football: England, France, Germany, Italy, Spain, the Netherlands

Generally, the results from these studies suggest that competitive balance differs across leagues, as is confirmed through the analyses in Chapter 4. The literature explaining differences in competitive balance across leagues focuses either on NAML, as for example Quirk and Fort (1992), or on the difference between NAML and European football leagues, as in for example Szymanski and Smith (2002). There has been little analysis of the determinants of the differences in competitive balance between European football leagues. The literature has focused on *comparisons* across a sample of leagues, rather than analysing the *determinants* of these differences. However, Bourg (2004) suggests differences in the distribution of revenues from sale of broadcasting rights as an explanation for different levels of competitive balance among the Big Five leagues.[149]

Differences in competitive balance between European football leagues and NAML have usually been explained in two ways. First, open leagues (i.e. merit hierarchies with promotion and relegation) in European football are seen to be advantageous for competitive balance in Europe in contrast to the closed leagues in Northern America. Second, the generally greater extent of regulations in NAML is seen as an advantage to these leagues, compared to European football. Note that the first point is related to win dispersion. Buzzacchi et al. (2003) and Szymanski and Smith (2002) find different results when it comes to performance persistence.

Differences in competitive balance between the NAML have, to a large degree, been explained by differences in regulations. However, when it comes to comparisons between the NHL and the NBA, Quirk and Fort (1992) apply a different approach, since differences in institutional rules are small and the teams have often used the same arenas. They therefore explain the differences in competitive balance between the NHL and the NBA on basis of "tumultuous" (p. 249), since the level of "expansion, contraction, and movement of league franchises" (p. 248) during the history of these leagues are different. Berri et al. (2005) focus upon a different approach, and follow Schmidt and Berri (2003) in applying the "Gould hypothesis". Schmidt and Berri (2003, p. 692-3) refer to the Gould hypothesis found in "the works of evolutionary biologist Stephen Jay Gould (1986; 1996) and economist Andrew Zimbalist (1992a; 1992b). Specifically, these authors argue that an expanding population of athletes would influence the convergence of team performance". Related to the NBA, Berri et al. find the lack of tall high quality players in the NBA as an explanation of the weak competitive balance in the NBA.

Fort and Quirk (1995) focus on revenue sharing policy when describing general differences in competitive balance between the NAML. A problem they explicitly emphasise is that within-league changes in the level of sharing have not been significant enough to allow empirical testing of the impact of changes in revenue sharing policy, related to possible changes in competitive balance. However, by comparing sharing

[149] Bourg (2004) analyses competitive balance over twenty seasons in the Big Five leagues in Europe and claims that (p. 5): "The position of each country can be explained by the more or less interdependent economic organization, operating mainly through the regulating function of broadcasting rights."

rules in the different leagues, they find that because of historic opposition to more equal sharing of gate and local TV-revenues in the NHL, MLB and NBA, these leagues have much less sharing than in the NFL, which historically has been much more open to such regulations. Further, Fort and Quirk (1995) claim that these differences might explain why the NFL has less disparity in "drawing potential" (p. 1289) than the other NAML. Hence, the revenues among the teams in NFL differ less than for the other leagues, as is confirmed by using Gini indexes. Fort and Quirk (1995) find that the index for the NFL is much lower than for the other leagues for the period 1991-1994, followed by the NHL, MLB, and finally the NBA as the weakest competitive balanced league. This is similar to predictions in Vrooman (1995), based on different institutional arrangements, such as payroll caps and factors affecting the revenue elasticity of winning (revenue sharing, the shared revenue as a proportion of all revenue, and the length of the season). Vrooman expects that the NFL should have best competitive balance, NBA least balanced, while the MLB should have moderate competitive balance.[150] These results are supported by Sandy et al. (2004, p. 178) claiming:

> ...in North America, larger market teams have dominated in baseball and basketball. Smaller market teams have had success in the National Football League, but this league shares the largest population of its wealth (Atkinson, Stanley, and Tschirhart, 1988; Mason, 1997).

5.3. Differences in Competitive Balance between European Football Leagues

Generally, most of the focus in the literature, when it comes to determinants for competitive balance, has been on two issues. The first is differences in drawing power (market size) among the teams in a league (see, for example, Fort, 2000; Gerrard, 2006b). The other significant part is related to possible effects on competitive balance from intervention on the free market, such as product and labour market restrictions.[151] This view follows the three points that, according to Szymanski (2003b), have been used in antitrust defences in U.S. courts (p. 1153):

[150] The NHL was not included in this analysis.
[151] These issues are consistent with "S2" and "S3" in Barros et al. (2002, p. 3), with the statement in Sandy et al. (2004) p. 178: "Virtually all leagues have debated the two primary tactics to ensure some level of competitive balance: (1) talent sharing; and (2) revenue-sharing".

1. Inequality of resources leads to unequal competition.
2. Fan interest declines when outcomes become less uncertain.
3. Specific redistribution mechanisms produce more outcome uncertainty.

In addition to these relationships, this thesis will also focus on possible incentive effects from prizes and league structure. These effects seem to be a minor part of the literature, even though a number of contributions have been related to relegation issues, in the context of competitive balance. In general, Gerrard (2006b, p. 34) claims that: "Economic theory demonstrates the possible impact of the corporate objectives of teams (i.e. wealth-maximisation or utility-maximisation), playing talent supply conditions, tournament structures including prize structures, and league regulations on the competitive balance of tournaments". These issues will be treated in the following, while, for example, the influence from technology and field conditions on competitive balance (Sanderson, 2002) is beyond the scope of this thesis.

The basic competitive balance course in professional team sports can next be explained by the following model:

Distribution of drawing power → Revenue distribution → Distribution of wages → Distribution of playing and coaching talent → Competitive balance

Given balanced match schedule, a league's competitive balance should fully reflect the distribution of sporting quality among the teams (sporting quality balance). The sporting quality balance is based on teams' ability to pay wages for the players. The stronger the correlation is between player (and coaching) quality and wages, the stronger the relationship between the distribution of wage costs and competitive balance should be, ceteris paribus.[152] The ability for teams to pay wages depends on their revenues. Hence, one expects a high correlation between the revenue balance among the teams in the league and competitive balance. Theoretically, perfect competitive balance should therefore be achieved if all teams were located in areas with equal revenue potential (El-Hodiri and Quirk, 1971). However, equal revenue potential is not easily determined in practice. A natural proxy might be the size of the population for a team.

[152] Kahane (2003) focuses on the balance of payroll when describing competitive balance. Even if Kahane lists a number of advantages for this view, this thesis regards it as a determinant (driver) for competitive balance.

Generally, population is found to be a strong significant demand driver in professional team sports (Cairns et al., 1986; Downward and Dawson, 2000; Dobson and Goddard, 2001; Borland and Macdonald, 2003).[153] However, differences in revenues among teams might also be related to other factors. Such factors can, according to Sanderson and Siegfried (2003), be differences between locations when it comes to the willingness to pay to have a successful team[154] and the preferences for winning.[155] Szymanski and Kuypers (1999, p. 263) claim that the "core level of support" of teams is related to both history and geography, where the latter is defined as the "catchment population and number of other clubs in the area". Empirical difficulties related to measuring teams' market size, or catchment area, are discussed in Dobson and Goddard (2001),[156] by referring to Dobson and Goddard (1992), and a number of factors are mentioned as indicators. These can be put into three main areas - demographic, socio-economic and historical, where the latter is related to historical records, traditions and long term sporting success. Especially for European football, it is important to include other factors than just population, because sporting success[157] and history of a team are significant contributors, and that the relative importance of geography might have been reduced compared to history, due to modern communications (Szymanski and Kuypers, 1999).[158] One would also, intuitively, expect the relationship between population and revenues to be weaker in European football leagues than in NAML, because the distance between competing teams are normally much smaller (see, for example, Dobson and Goddard, 2004). Summarized, the revenue distribution among the teams in a European football league is a function of their *drawing power* (a term applied by, for example, Hoehn and Szymanski, 1999; Dobson and Goddard, 2001), and is, hence, the basic determinant for a league's level of competitive balance.

[153] Another advantage might intuitively be that teams from larger populations should have higher probability of being demanded by high level talents, who would prefer to play for "their" team, if expecting a positive correlation between population and the number of high level talented players. This can be both monetary in the sense of reduced living costs, but also immaterial variables such as playing for the team the player is supporting, and living in the area where the players' friends live. Analyses of these immaterial advantages are beyond the scope of this thesis.

[154] According to Sanderson and Siegfried (2003) this willingness might be dependent on the alternative opportunities on recreation, entertainment and/or cultural amenities.

[155] Sanderson and Siegfried (2003) use Porter (1992) as reference.

[156] See also Dobson and Goddard (2004).

[157] See, for example, Cairns (1987).

[158] The former coach of Rosenborg BK claims that Rosenborg BK, when qualifying to the top division in the Norwegian league for the first time for the 1960/61 season, became popular among others because of the style of the team, young up and coming players playing attractive offensive football (Eggen and Nyrønning, 1999).

During the history of professional team sports, cross-subsidisation methods have been popular tools, aiming to weaken the relationship between the distribution of drawing power and the level of competitive balance. Among the motives for introducing and keeping these regulations is the uncertainty of outcome hypothesis. In other words, through pro competitive balance cross-subsidisation policy, one expects increased demand through the uncertainty of outcome hypothesis, ceteris paribus. Sloane (1971) claims that this may "justify" the sports leagues' use of it, even if cross-subsidisation in general is criticised by economists (see also Jennett, 1984). According to Downward and Dawson (2000), one can, therefore, say that the uncertainty of outcome hypothesis promotes cross-subsidisation policies. However, to achieve the proposed effects on uncertainty of outcome, it is required that cross-subsidisation policy has positive effects on competitive balance. The literature of the economics of professional team sports has concerned effects from cross-subsidisation policy both theoretically and empirically. This chapter will only focus on competitive balance effects.[159]

The main focus on regulations in a competitive balance context is related both to policy affecting the distribution of revenues among the teams and on restrictions on the labour market. Revenue sharing policy (regulations on the product market) is mainly divided into two categories; the traditional gate revenue sharing policy and the modern media revenue sharing policy. Regulations on the labour market are related to the player transfers, wages, player recruitments, and closed labour market. Since many of the labour market regulations are similar between the leagues in "modern" European football, they will not be included in the model for determinants for differences in competitive balance in European football.

On the product market, revenue sharing has a long history in professional team sports. Motivations for justifying a revenue sharing system might be divided into two parts; reduced financial divergences among the teams and what this thesis will call "reasonableness". The latter refers back to the peculiarities of team sports, and that two teams are required to produce a match (see, for example, Rottenberg, 1956; Neale, 1964). A broader revenue sharing system among the teams in a league can also be

[159] Other effects might be a general financial benefit for the weaker teams both in relation to money redistribution (from wealthier teams) and as a consequence of reduced bargaining power for players (benefits all teams). This can for example increase the probability that they will stay in business (Dobson and Goddard, 2001). These arguments refer back to the starting period of both Major League Baseball (Rottenberg, 1956) and English football (Szymanski and Kuypers, 1999).

justified, because the league in itself might be a main revenue driver for *all* teams (Gerrard, 2006b).

Already at the beginning of the National League in 1876, the ticket price was equally shared between the home and away team (Fort and Quirk, 1995, p. 1286). According to Barros et al. (2002), gate (and merchandise) revenue sharing has played a minor part of the economics of European football, even though it has been a part of the history of for example English and Norwegian football. However, both these leagues have removed the gate sharing between the home and away teams. This chapter will not concern gate sharing as a significant determinant for differences in competitive balance between European football leagues.

Historically, revenues from broadcasting deals were limited, both as a consequence of weak competition among TV suppliers, which often were public broadcasters in a monopoly situation (see, for example, Barros et al., 2002),[160] and that the fear of substitution effects (potential spectators at the stadium are watching matches on television) among governing bodies reduced the interest to televise these matches. Deregulation in the market for broadcasting in Europe during the 1980s (see, for example, Jeanrenaud and Késenne, 2006) seems to have created a significant (positive) shift in the demand for broadcasting rights. In addition, the consequential expanded level of compensation for these rights has (probably) further increased the interest from governing bodies to sell the product to broadcasters, since possible substitution effects now could be (over-) compensated.[161] Moreover, positive spill-over effects from television on merchandising, sponsors and promotion (in a marketing context) of teams and the sport have forced changed behaviour among leaders of teams and governing bodies. Especially from the 1990s, innovative infrastructure among broadcasters have created a further positive shift in the demand curve for broadcasting rights, and hence inflated prices even more.[162]

Solberg (2004) defines sports rights as (p. 378): "the rights to broadcast from a sporting event, most commonly within specific geographical areas." According to Solberg, the

---

[160] See for example Barros et al. (2002) for discussions about these issues. They also refer to a number of relevant studies.
[161] This is described in the analyses of the effects on demand from the second (1993/94) season of the first deal between the BSkyB and the English FA Premier League by Baimbridge et al. (1996).
[162] Solberg (2004 p. 382-6) and Gratton and Solberg (2004, 2007) provide overview over the European TV landscape.

sports rights can further be split into different groups of products, such as live broadcasting and highlights. In addition, for example football rights have been a driver for developing pay television in Europe (Jeanrenaud and Késenne, 2006). The methods for constructing and organising the deals differ substantially between the members of the UEFA. This means both the framework of the deals (individual in contrast to collective deals) and the general contents of the contracts (e.g. the percent to be equally distributed). The rights are sold in three general ways; collective sale,[163] individualistic sale, and combinations of these two. Collective sale of broadcasting rights is still most common and is now the major revenue sharing source in many European football leagues.[164]

With collective sale of broadcasting rights, the revenues are distributed among the teams in a way that is expected to be narrower than for individualistic sales. This is emphasised in Jeanrenaud and Késenne (2006), claiming that the individual deals are an advantage for the big-market teams, and hence lead to a weaker distribution of broadcasting revenues among the teams.[165] Empirically, this can be shown in for example Bourg (2004), who analyses the broadcasting deals in all Big Five leagues and claims that the differences in constructions have expanded the domination among the dominating teams in Italy and Spain. This is emphasised by calculating the ratio for differences in the earnings for the team receiving the highest amount and the club receiving the lowest amount of money from broadcasting. The ratios (ranges) are "1.7 in France, 2.2 in England, 2.6 in Germany, 5.3 in Spain, and 6.3 in Italy" (p. 9).[166] This is consistent with the findings of Solberg (2004), who compares the differences between the collective deal in the English FA Premier League and the individualistic deal in Italian Serie A. However, according to Jeanrenaud and Késenne (2006), Tonazzi (2003) emphasises that individual broadcasting deals not necessary weaken competitive balance, if an appropriate revenue redistribution mechanism from bigger to smaller teams exists.

[163] Jeanrenaud and Késenne (2006) use the expression "joint selling". They refer to a number of advantages of this type of selling the broadcasting rights, such as lower transaction costs and increased possibility to buy a complete package for a whole season.
[164] Downward and Dawson (2000, p. 169) claim that: "By implication, the skewed nature of television revenues in European sport would seem to suggest that a lack of competitive balance, which historically seems to be the case, would be either, at a minimum, consolidated or, at a maximum, accelerated".
[165] Jeanrenaud and Késenne (2006) use Scully (2004) as reference.
[166] The ratio for England is consistent with Sandy et al. (2004).

In NAML, the history of significant revenues from sale of broadcasting rights is much longer than in Europe. In general, the national rights are distributed uniformly among all teams (some exceptions for expansions teams), while the way to treat local TV-revenues vary between the leagues. Theoretical analyses of competitive balance consequences from broadcasting revenues have been split into two in NAML. Equal sharing of national TV revenues should not have any effect on competitive balance, since the payment is not related to the teams' playing success (Fort and Quirk, 1995), while sharing of local TV revenues might have positive effects on competitive balance (Fort and Quirk, 1995). However, the deals when it comes to sale of broadcasting rights differ across the Atlantic, and these theoretical results might not be directly transferred to European football. In a number of European football leagues, parts of the distribution from collective revenues from sale of media rights are based on sporting performance.[167] According to Szymanski (2001, 2003c), this should promote competitive balance.

Professional leagues across countries and sports both have similarities and differences in relation to prizes and tournament structures. A number of differences can be found in relation to tournament structure and organization between the NAML and European football leagues (Hoehn and Szymanski, 1999).[168] For example, the major leagues are closed, while the European football leagues are generally open. The effects of prize and tournament structures on competitive balance have generally not received much attention, at least not in European football (Barros et al., 2002). However, there have been analyses of competitive balance effects from promotion-relegation (Hoehn and Szymanski, 1999; Noll, 2002, 2003), expansions (Fort and Quirk, 1995; Schmidt, 2001), reduced number of teams (Cairns, 1987), post-seasonal play (Sandy et al., 2004; Bourg, 2004; Goossens and Késenne, 2007) and point score systems (Cain and Haddock, 2006; Fort, 2007; Haugen, 2008). In addition, Lee and Fort (2005) and Fort and Lee (2007) among others use different structural changes in analyses of their break-point analyses of the NAML. Utt and Fort (2002) recognise that the playing schedule system in the MLB might affect competitive balance in the description of perfect competitive dominance, while Ross and Lucke (1997) explain changes in competitive balance also from changing to divisional play in the MLB. Szymanski (2003b) discusses

[167] According to Sandy et al. (2004) the only performance based revenues is from sale of broadcasting rights related to qualifications for playoffs.
[168] An overview over "common" elements in the NAML and European football can be found in Szymanski (2003b, p. 1150-2).

optimal size of leagues, and claims that this has been a relevant issue in the major leagues, while in European football it is an issue that "have never risen" (p. 161).

5.3.1. Model

Based on the discussions above, analyses of determinants for differences in competitive balance between leagues in European football require a different approach than for NAML. This is based on three major groups of differences between the leagues on each side of the Atlantic. The first group of differences is related to non-sporting variables, such as demographic and economic factors. The differences in these variables are relatively small when comparing the NAML, because they operate in the same area. On the other hand, there are significant variations among these variables at the macro level between the members of the UEFA. The second group is related to market regulations in the leagues. In general, the differences in competitive balance between the NAML have, to a large extent, been explained by significant variations in regulations between the leagues. In Europe, the situation is quite different, since the differences in the level of market regulations are very limited between the different members of the UEFA. Therefore, market regulations provide limited explanation of the differences in competitive balance between European football leagues, with the possible exception of differences in the distribution of the money earned by sale of broadcasting rights, as indicated in Bourg (2004). The third group is related to prize and tournament structure. While these are relatively similar across the NAML, there seems to be larger divergences among European football leagues.

Analysing the determinants for differences in competitive balance between the different members of the UEFA will be done by first analysing possible determinants for dispersions of drawing power from a macro level. This will be done through including explanatory variables for the distribution of drawing power among the teams. Next, the structure of revenue sharing through TV/media deals will be analysed, while the last group of variables will concern prize and tournament structure.[169]

The framework of the cross-sectional analysis is based on the following model:

---

[169] Table 2 in Bourg (2004) p. 8 summarizes convergences and divergences in the organization of professional team sports in Europe.

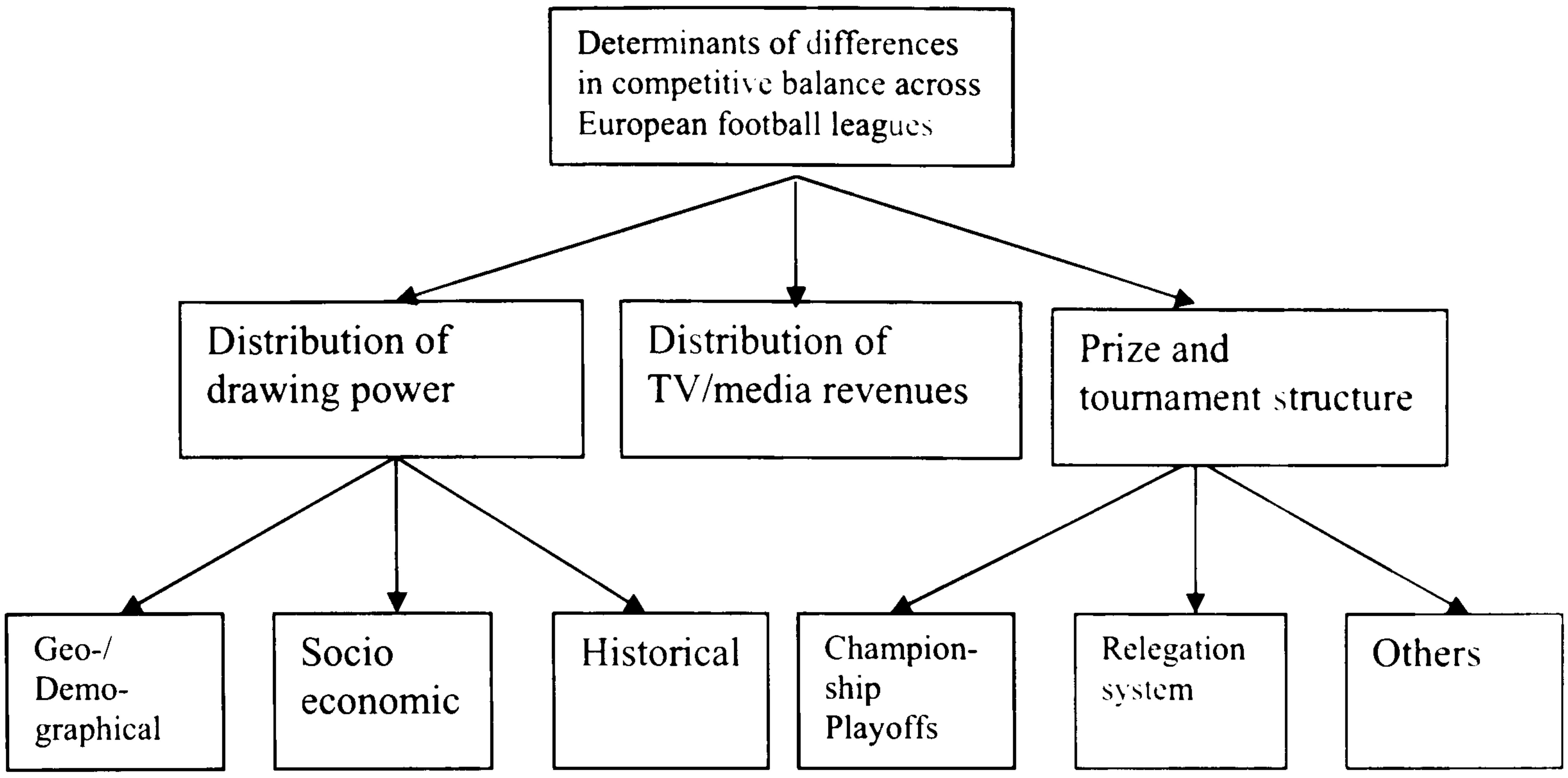

Figure 5.1: Overview over factors affecting competitive balance between European football leagues

Geographic/Demographic Variables[170]

Many studies claim that the most important determinant in a team's drawing power is the population in the "supporting zone" for the team. Because this is an analysis on the macro level, general variables are applied in comparison to within-league analyses. Two issues will be focused on in this part; market size and the distribution of population within the different leagues (countries). The first is based on the assumption that a bigger market size for a sport increases the distribution of drawing power among the teams, ceteris paribus. The second is based on the fact that fans have high tendency to support teams from their home area, and the following hypothesis can be drawn: the greater distribution of the population there is throughout the country, the better competitive balance there will be. Moreover, population might be a reflector for the Gould hypothesis, where a larger population can be hypothesised to bring a higher number of talented players to share among the teams, ceteris paribus. However, this possible relationship between population and competitive balance might be much weaker after the Bosman ruling, due to the open labour market in the EU.

The obvious demographic variable to use for total market size is population. For example, Lee (2004) uses population as a measure of market size, when comparing

[170] Literature based on urban studies could be relevant for this group of variables, but it is seen as beyond the scope of this thesis.

attendance in different leagues. In this chapter's cross-sectional model, the total population (POP) for the country of the league is used. In addition, there are variables that might be more specific in reflecting the distribution of population within a league. In other words, better competitive balance is expected in countries with a more equal spatial distribution of the market. Relevant variables are: population density (POPDENS), population of the largest city (LARGCITY), population of the largest city as a share of the total population, measured by dummy variables (LARGSHARE > 0.25; SMALLSHARE < 0.1),[171] and percentage size of the agricultural sector (AGRI). Population of the largest city as a share of the total population might reflect that the population is concentrated around this city, and hence predict a lower competitive balance. The size of the agricultural sector will also tend to reflect population density (with more agrarian-based economies tending to have lower population density). High percentage of agriculture might therefore indicate more equal distribution of the population within a league.

Economic Variables

The next group of variables is economic variables. More highly economically developed countries are hypothesized to have more equal spatial distribution of the market. In addition, this variable might also be positively correlated with determining the market size of a league. Lee (2004), when comparing attendance between NL and AL in MLB, the Central League (CL) and the Pacific League (PL) in the Japanese Professional Baseball League (JPBL), and the Korean Professional Baseball League (KPBL), uses a variable based on per capita GDP. I will use the per capita GNP as a variable (GNP) to reflect the economic size of the country. The percent of agriculture (AGRI) might also reflect economic development, and will, in this case, have the opposite expected sign, compared to the population explanation above.

Sporting (Historical) Variables

Sporting variables may reflect at least two issues, hypothesised to be relevant in the context of competitive balance. First, the general sporting quality of a league is expected to be positively correlated with competitive balance. Second, the general

[171] The split into dummy variables fits better than just applying a continuous variable.

market size for a league is not only reflected by the population and, eventually, economic conditions, but also from historical sporting traditions.

The sporting quality of a league in European football can be measured by comparing the sporting level of the teams of the league. The current system of European team tournaments, such as the UEFA Champions League, includes a seeding system that is league based, and this is applied as a measure for the level of sporting qualify in the league (UEFARANK).

The market size of a league is also determined by how popular the sport is. It is expected that more popular sports (measured by spectator demand) will tend, ceteris paribus, to generate greater supply of playing talent. The variable used is the FIFA rank system for national teams (FIFARANK). The age of the league might also reflect the popularity of the sport, but this can also be affected by political changes. However, the age of the league can be relevant in the context of competitive balance, if it is expected that developed leagues are better balanced, ceteris paribus. This hypothesis is supported by the findings in Lee (2004), who concludes that there is a higher degree of uncertainty of match and playoff outcome in developed leagues than in developing leagues. Because of the structure of observations for this variable, a dummy variable is preferred to identify newer members of the UEFA (UEFAlate). An association becoming member in 1988 or later is defined as young.

Revenue Sharing

The model for analysing determinants for differences in competitive balance across leagues in European football also includes revenue sharing. However, the relevant variables for revenue sharing in these leagues are based on broadcasting deals. Because of lack of information, variables for broadcasting rights will be treated in a second step of the model, in residual analyses in the Big Five leagues.

Prize and Tournament Structure

Additional to the macro level variables suggested to proxy leagues' distribution of drawing power, the focus on prizes and tournament structure in this chapter gives the opportunity to include variables at the micro (single league) level. It might be reasonable to expect that aspects of tournament structure, such as championship playoffs and the relegation system, might affect the level of competitive balance in a league.

The challenge is to include appropriate variables into the competitive balance equation. For championship playoffs, a dummy variable (ChPO) takes the value of one, if the league has had championship playoff during the periods analysed. For relegation, the difference between leagues is related to the number of teams that are involved in relegation systems, either through automatic relegation or through relegation playoff. Two variables will be used in this analysis. The first is the number of teams to automatically relegate in relation to the number of teams in the league (AuRel_nr). The hypothesis is that competitive balance is expected to improve, as the higher share of teams to be automatically relegated increases. Second, since many of the European football leagues also include relegation playoff in their relegation procedure, a variable that uses the total number of teams involved in the relegation process as share of the total number of teams in the league (AuPoRel_nr), is included. The hypothesis is based on the same as the previous variable.

Most European football leagues use what can be called a pure round-robin match schedule, where all teams meet each other twice, once at home and once away. Deviation from this system might affect competitive balance, if weaker teams get advantages from the match schedule, compared to better teams. The dummy variable (PureRR) has the value one for leagues using pure round-robin, and zero otherwise.

The last variable is related to the size of the league by the number of teams (AVteams). Even if the measures for win dispersion and performance persistence take into account the number of teams in the league, the size of the league might have further effects on competitive balance. Ceteris paribus, one might expect that the more teams there are in the league, the weaker the competitive balance, because weaker teams are expected to be included as the league increases its number of teams. This should worsen the win dispersion.

As is shown earlier, point score systems might affect some of the competitive balance measures, in addition to possible effects on competitive balance through sporting incentive effects. Except for Albania in the 1995/96 season, who only had two points for a home win (but three for an away win), all member leagues of UEFA operate with the (3,1,0) point score system during the sample period. No difference between the leagues should emerge because of point score system in the cross-sectional analysis presented below, and this is therefore omitted from the model.[172] On basis of the discussion above, the model for analysing the determinants for differences in competitive balance across leagues in European football is shown in the following equation:

$$CB_i = \beta_1 + \beta_2 POP_i + \beta_3 POPDENS_i + \beta_4 LARGCITY_i + \beta_5 LARGSHARE_i + \beta_6 SMALLSHARE_i + \beta_7 AGRI_i + \beta_8 GNP_i + \beta_9 UEFARANK_i + \beta_{10} FIFARANK_i + \beta_{11} UEFAlate_i + \beta_{12} ChPO_i + \beta_{13} AuRel\_nr_i + \beta_{14} AuPoRel\_nr_i + \beta_{15} PureRR_i + \beta_{16} AVteams_i + \varepsilon_i$$

5.3.2. Data and Descriptive Analyses

The data in this study is collected from a number of sources. Economic and geographic data are found in the World Factbook,[173] McCoy (2002), www.world-gazetteer.com, and www.nationmaster.com/graph-T/geo_pop_den/EUR. Sources on sporting data are www.uefa.com, rsssf.com, www.uefa.com/uefa/keytopics/kind=64/newsid=38404.html and www.fifa.com/worldfootball/ranking/. Competitive balance is calculated annually for 50 of the 52 members of UEFA in two periods (five and ten seasons), within the following sample period of ten seasons: 1994(/95)-2004(/05).[174] The five season analysis is based on the last part of this period; 1999(/00)-2004(/05). The variables are presented in table 5.23 and 5.24 in the appendix of this chapter.

Competitive balance is calculated for the three dimensions used throughout the thesis. The NSQF ratio, SRCC and HHI are calculated for all leagues. A difficulty is the choice of appropriate timeframe for these variables, because in general, within-seasonal competitive balance varies considerably between seasons. Cross-season averages are

[172] However, effects from changes in point score systems on competitive balance are possible in the time-series analyses in Chapter 6. Also differences in corporate objectives is difficult to differ between leagues, but will be discussed in Chapter 6 too.
[173] www.odci.gov/cia/publications/factbook/index.html.
[174] As the only UEFA member, Liechtenstein does not have its own league, and can hence not be included in the data sample (teams from Lichtenstein play in the Swiss league). In addition, the structure of the league in Bosnia-Herzegovina complicates calculations of competitive balance, and is also omitted from some of the analyses in this chapter.

used to eliminate random fluctuations over time. Also many of the independent variables used in this analysis, such as population, are relatively stable in the short run. The number of seasons is arbitrary, and is for the NSQF ratio and the SRCC on ten and five seasons, while the HHI is only based on ten seasons.

For general comments about competitive balance in European football, see Chapter 4. However, one aspect is relatively clear, that among the weakest balanced leagues (NSQF ratio) are young (and smaller) leagues from the former Soviet Union. Among the members of the UEFA in this analysis, 27 leagues have been members from the foundation of the association; while as many as 22 of today's members in the data sample entered the organisation after 1987.[175] In other words, the data might give indications of great divergence in age, and hence, the history and tradition, of different European football leagues.

The member leagues of the UEFA have a wide dispersion of all kinds of variables at macro level, since they in general (but not all) are independent nations. One example of the wide disparity in a variable at macro level is the population for each league. The extremes are about 144 millions citizens in Russia, while the population in San Marino is less than 30,000 citizens. The mean population for a league in European football is 16.37 million citizens. However, it is obvious from the data that the population in a number of leagues is relatively high compared to the rest. Ten of the leagues have a population above 20 million citizens, while 19 of the leagues have a population above ten million citizens. Since as many as 32 out of the 51 leagues (including Bosnia-Herzegovina) have a population that is less than ten million citizens, the median might give more appropriate information of the typical population in an average European football league; 5.42 million citizens. This means that most of the leagues in European football are relatively small, but at the same time, there are a number of leagues that are much bigger than the others.

Also for the population density, variations across European countries are wide, from Iceland as the lowest, with 2.72, to Malta, with 1,192.51 people per squared kilometre. Among the bigger leagues, the Netherlands have a population density of 466.45 and England of 380.40, while Russia has 8.61 and Spain has 78.43 people per square kilometre. The population of the biggest city relative to the total population of the

[175] Including Liechtenstein.

country varies from Iceland (0.53), Armenia (0.44) and Latvia (0.40) to Ukraine (0.06), Poland (0.04) and Germany (0.04).

In addition, the agriculture percent differs considerably between the leagues, from Albania having a percentage of 47.5 to Luxemburg with only 0.5 percent in the agriculture sector.[176] Differences in economic development can be shown using a variable of GNP per capita. Also this variable varies significantly between the different leagues in the UEFA area, from Luxemburg having $ 55,100 to Moldova having $ 1,800.

Revenue Sharing in European Football - Broadcasting Rights

At least two measurement problems are related to broadcasting revenues in the cross-sectional analysis across European football leagues. These are changed distribution during the periods to be analysed and lack of information in a significant number of leagues.[177]

Differences in broadcasting revenues across the teams in a league might be based on a number of sources. A main difference between leagues, as well as within some leagues over time, is the structure of the broadcasting deals; are deals collective and/or individual? According to Andreff and Bourg (2006), the top division in Greece, Italy, Portugal and Spain[178] operated with individual deals in the 2002/03 season. Andreff and Bourg calculate that the Gini coefficient for the TV rights in Serie A in Italy increased from 0.211 to 0.430, between 1998/99 and 1999/00, as the league went from collective to individual deals. Note that the calculation of 0.430 is related to the non-pooled share of the rights.

There can also be general differences when it comes to the share of the deals that are equally distributed. Andreff and Bourg show that the "solidarity share" in the French league went from 91 percent in 1998/99 to an expected 50 percent in 2003/04. Another

[176] Because of missing information of this variable in San Marino and Andorra, this chapter will use arbitrary percentages of agriculture percent in these leagues.
[177] It might be that this part should have been named even broader, because there are examples that it is not only broadcasting rights that are shared between teams in a league. Internet and other media revenues might also be included in the packages that are shared between teams.
[178] Broadcasting deals in Spain are based on a mixed system, as they are both individual and collective (Gratton and Solberg, 2007). More information can be found in TV Sports Markets (2006).

source for differences in the distribution of broadcasting revenues can for example be shown in table 3.2 in Andreff and Bourg (2006), where they split between broadcasting rights in relation to *free to air, pay-TV, pay-per-view and international* rights. The method for distribution of these revenues may differ between the different parts of the broadcasting rights.

Differences are also related to both the absolute level of broadcasting revenues[179] and from these revenues' share of the total revenue for different teams in different leagues (Andreff and Bourg, 2006), and to differences in sports between countries (Jeanrenaud and Késenne, 2006).

Descriptive data for residual analyses, related to the distribution of revenues from broadcasting rights in the second part of the cross-sectional analysis, is based on the calculations of the range (as a ratio) between the team with the highest level of these revenues, compared to the team with the lowest amount of broadcasting revenues (Andreff and Bourg, 2006).

Table 5.2: Range ratios of broadcasting revenues from table 3.6 in Andreff and Bourg (2006)

Index	Season	Spain	Italy	Germany	England	France
Range ratio	2000/01	5.3	6.3	2.6	2.3	1.8

These results seem to have relatively high correlation with the level of redistribution/solidarity percent among these leagues.

Number of Teams

As mentioned before, there are a number of differences in the organisation of the leagues in European football and NAML. This is also related to league structure, both with regards to the number of teams and the number of divisions at horizontal level. In European football, the number of groups in the top tier is usually one. All UEFA member leagues, except from San Marino (where the top division is split into two groups with seven teams in Group A and eight teams in group B), had this structure in the 2004(/05) season. However, the size of the top division varies across these leagues, as shown in table 5.3.

179 Gerrard (2006b) provides a discussion of the main factors affecting the media value of team tournaments.

Table 5.3: The number of teams in top divisions in European football in 2004(05)

Number of teams	League
20	England, France, Italy, Spain
19	Kazakhstan
18	Azerbaijan, Belgium, Germany, The Netherlands, Portugal, Turkey, Wales
16	Bosnia-Herzegovina, Belarus, Bulgaria, Croatia, Czech Republic, Greece, Hungary, Northern Ireland, Romania, Russia, Serbia-Montenegro, Ukraine
14	Cyprus, Finland, Norway, Poland, Sweden
12	Denmark, F.Y.R. Macedonia, Israel, Luxembourg, Scotland, Slovenia
10	Albania, Austria, Faroe Islands, Georgia, Iceland, Malta, Republic of Ireland, Slovakia, Switzerland
8	Andorra, Armenia, Estonia, Latvia, Lithuania, Moldova

In the cross-sectional regression analysis later in this chapter, the number of teams during the analysis period (ten or five seasons) is averaged for each league.

NAML are organised differently, compared to European football leagues, both with regards to the number of teams and the number of (horizontal) groups at top level. These leagues have under groups (conferences) that are again split into divisions, which are all at the same horizontal level. The MLB, NBA and the NHL consist of 30 teams. These leagues are split into two conferences of three divisions. Each division in the NBA and the NHL consists of five teams (= 15 teams in each conference). MLB does not have the same symmetry since one division has six and another has four, additional to the four with five teams. This means that one conference (American League – AL) has 16 teams, and the other (National League – NL) has 14 teams. The NFL has 32 teams, as are divided into two conferences of 16 teams each from four divisions of four teams each.

In summary, the top division in European football leagues is much smaller than the NAML. Even though the number of teams is higher in NAML, the market size for each team is in general much higher. Sandy et al. (2004) claim that this might be one of the reasons for the generally weaker financial position of European teams.

Relegation Systems

In the 2004(/05) season, all member leagues of the UEFA operated with relegation systems, except for the closed league in San Marino. There are three procedures in European football when it comes to relegation system. First, there is automatic relegation. By automatic relegation is meant that the team/teams at the bottom of the end-of-season regular season is/are relegated to the next (lower) level division without any options. The second system is relegation through a two-step procedure of "qualifying" for relegation playoff, and with the poorest performing team (s) in this playoff being relegated. The third system is a combination of these two methods. Table 5.4 shows the number of teams to be automatically relegated in European football leagues that have only automatic relegation in the 2004(/05) season.

Table 5.4: Overview over the European football leagues having only automatic relegation in the 2004(/05) season

1 team	Armenia,[1] Austria, Republic of Ireland, Slovakia
2 teams	Albania, Belgium, Bosnia-Herzegovina, Czech Republic, Denmark, Georgia, Hungary, Iceland, Israel, Russia
3 teams	Bulgaria, Cyprus, England, France, Germany, Greece, Portugal, Romania, Serbia-Montenegro, Spain, Turkey, Ukraine, Wales
4 teams	Azerbaijan
Total	28 leagues

[1] After the 2004 season it was decided not to relegate the team due to expansion.

The other usual relegation system is a combination of automatic relegation and relegation playoff. The organization of the number of teams to automatically relegate and the number of teams to relegation playoff vary across the UEFA member leagues in the 2004(/05) season. An overview is provided in the table below.

Table 5.5: European football leagues having combined relegation systems in the 2004(/05) season

1 team auto + 1 team playoff	Estonia,[1] Faroe Island, Finland, Latvia, Moldova, Northern Ireland, Poland, Switzerland
1 team auto + 2 teams playoff	Belarus[2]
2 teams auto + 1 team playoff	Norway, Sweden
2 teams auto + 2 teams playoff	F.Y.R. Macedonia, Georgia[3]
3 teams auto + 1 team playoff	Kazakhstan

[1] No automatic relegation in 2004 due to expansion.
[2] The two teams qualifying for playoff played against each other.
[3] This is based on the 2003/04 season, because the next season had no playoff as a consequence of two teams being excluded prior to its start.

Among the top divisions in European football leagues, relegation systems only based on playoff (no automatic relegation), either include second level division teams or only consist of top division teams. For both systems, the starting point score in the playoff differs between leagues. Some have the starting point score independent (zero points) of regular season, while others have the start position dependent on regular season performance.

For relegation playoff *only consisting of teams from the top division;* no European football leagues arranged this system where the *starting point was independent on regular season performance* (ignoring eventually match schedule influence) in the 2004(/05) season. However, Azerbaijan used this method in the season 1998/99, when four teams went into playoff, and two of them were relegated.[180] Most of the leagues using relegation playoffs that only include top division teams have a starting point that is dependent on performance in regular season. The table below shows these leagues for the 2004(/05) season.

Table 5.6: European football leagues having a relegation system only based on top division teams in the 2004(/05) season

Number of teams into playoff	Number of teams to relegate	Points from regular season	League
8	2	100 %	Luxemburg[1]
6	1	100 %	Scotland
6	1 + 1	100 %	Slovenia[2]
4	2	50 %	Malta

[1] The bottom team from each of the two relegation playoff groups was relegated.
[2] This is in principle. However, in practice, three teams were relegated due to financial problems this season. Consequently, the second least team in the relegation playoff escaped from a new relegation playoff (against the runners-up in the second level division) and the league was reduced from twelve to ten teams.

As indicated in Chapter 3, this thesis defines the Scottish Premier League as a league with a regular season (with unbalanced match schedule, see table 5.11 below) and post-seasonal playoff for both the championship and for relegation. The Scottish league is therefore included in the table above.

The other general way to arrange relegation playoff (without automatic relegation) is to also *include the better teams from the second level division(s).* Through this kind of

[180] Anticipating that the bottom four teams were going to the relegation playoff, and that bottom two in the relegation playoff should relegate. However, one of the teams (number 13 of 14) withdrew after the first stage, and did not participate in the playoff. Therefore only three teams participated, and of these three teams only one relegated.

playoff, the best teams will win participation in the next season's top division, while the other teams will be members of the lower level division(s). This kind of system, with *independent starting point from performance in the regular season*, was used in one European football league for the 2004(/05) season, the Netherlands. The bottom two teams from the top division, together with six teams from the second level divisions, competed for two places in the Eredivisie (top division). No leagues in the 2004(/05) season used a structure in which the *starting point is not independent from regular season*. However, Finland came into this category in the 2002 season, when four teams from the top division and two teams from each of the second level divisions met to play for six places in the top division (expansion). The starting point was based on a bonus point system, where number nine in the top division got three points, number ten got two points, number 11 one point and number 12 zero points. In addition, the winners of the second level divisions also had one bonus point, while the runners-up had zero points.

In the cross-sectional regression analysis later in this chapter, the number of teams in each category is averaged over the analysis periods.

In Latin-America a number of leagues operate with a different method in deciding which teams to relegate, compared to European football leagues. Relegation is decided on performance, not only from the current season, but also from earlier seasons, in leagues, such as in Argentina (2004/05), Columbia (2004) and Mexico (2004/05). Analysing effects from these relegation systems are beyond the scope of this thesis.

As can be seen from the overview above, there is a number of ways to structure the movements between the tiers in open leagues. The relegation system in leagues can also be used in relation to changing the number of teams participating in the top division. Note that teams can also be relegated as a consequence of administrative penalties, which might have been more relevant in European football, after the introduction of the so-called UEFA licence.[181]

[181] The Norwegian football league for women has a club licence system, and after the 2005 season a team (Asker FK) was relegated for not justifying certain requirements for passing through the licence system. In the Italian league, Juventus was relegated after the 2005/06 season as an administrative penalty. More about the UEFA licence can be found at www.uefa.com, and more specific for the UFEA club licence system by 2005: www.uefa.com/newsfiles/358508.pdf.

Championship Playoffs

Another main difference in league structure is related to the way the championship winner is decided. More precisely, does a certain league have post-season championship playoff, or is the championship winner decided only on basis of regular season play? The existence of championship playoff in European football differs between leagues. Some of them, such as the English top division, have never had the championship winner decided after playoff, while it is a part of the current system across the border, in the Scottish league.

The structure of championship playoff diverges, both between leagues and within leagues over time. In general, there are three main differences in the structure across leagues. These are differences in how the championship playoff tournament is played, the number of teams to championship playoff, and the relationship between the regular season and the playoff. The latter might vary from total independency to full transfer of the regular season score. Bonus points at the start of the playoff is another way to value regular season differences. In addition, some leagues operate with seeding systems on basis of regular season performance.

Below, an overview is presented, showing the European football leagues that have used some kind of championship playoff in its leagues structure during the period 1995(/96) – 2004(/05).[182] This is also the basis for categorisation of the championship playoff dummy variable, applied in the cross-sectional regression analysis. The table below consists of the 26 out of 51 European football leagues that had at least one season with championship playoffs during this period. This table also includes information of the type of championship playoff that is played, if it has league or knock-out format.

[182] This limitation might "hide" that some leagues have had restructurings close to this period, such as the Irish league that had (the last) championship playoff in the 1993/94 season, and will hence not be treated as a league with championship playoff in the 1995(/96)-2004(/05) period.

Table 5.7: Overview over European football leagues having a championship playoff system over the period 1995(/96) – 2004(/05)

League	**1995 (/96)**	**1996 (/97)**	**1997 (/98)**	**1998 (/99)**	**1999 (/00)**	**2000 (/01)**	**2001 (/02)**	**2002 (/03)**	**2003 (/04)**	**2004 (/05)**
Andorra						L	L	L	L	L
Armenia				L						
Azerbaijan	L			L			L			
Bosnia-H	ko	ko	L	(--)	L					
Bulgaria							L			
Croatia	L		L	L		L				
Estonia	L	L	L							
Finland		L			L			L		
Georgia					L	L	L	L	L	
Hungary						L	L	L	L	
Israel						L				
Kazakhstan								L		
Latvia	L	L								
Lithuania	L									
Luxemburg					L	L	L	L	L	L
Macedonia							L			
Malta					L	L	L	L	L	L
Moldavia				L						
N. Ireland								L		
Poland							L			
San Marino	ko	ko	ko	ko	ko	ko	ko	ko	ko	ko
Scotland						L	L	L	L	L
Serbia Mont	L									
Slovakia	L									
Slovenia									L	L
Switzerland	L	L	L	L	L	L	L	L		
Sum	**10**	**6**	**5**	**6(--)**	**7**	**10**	**12**	**11**	**8**	**6**

L = League system (might include the final as a knock out game).
ko = Knock out system.
(--) = Cancelled playoff.

While all NAML have annual post-seasonal championship playoffs, these kinds of playoffs have only been a feature of some of the European football leagues, and then often only for a limited period of time. Only one league in European football operated with a championship playoff during the whole sample period - San Marino. Bosnia-Herzegovina also had a championship playoff with a knock out tournament system in the 1995/96 and 1996/97 seasons. All other championship playoffs in European football in the period 1995(/96)-2004(/05) are league based. In general, the playoff structure in European football is therefore different from the NAML, where playoffs are arranged as knock-out tournaments.

The second difference in the championship playoff system is the number of teams that qualify for championship playoff. In the actual sample period, the number of teams to qualify for post-seasonal championship playoff varies among the European football leagues. This is shown in the table below.

Table 5.8: The number of teams to championship playoffs in European football leagues over the period 1995(/96) – 2004(/05)

Number of teams to championship playoff from regular season	League (seasons with deviating structure in brackets)
4 teams (one regular season group)	Andorra, Luxembourg and San Marino (+ the winner of the second level division – 1995/96)
5 teams	Moldova and Croatia (1995/96)
6 teams	Armenia, Azerbaijan, Bulgaria, Croatia, Estonia, F.Y.R. Macedonia, Finland (1996), Georgia, Hungary, Israel, Kazakhstan, Latvia, Malta, Northern Ireland, Scotland, Serbia-Montenegro (Yugoslavia), Slovakia and Slovenia
8 teams	Finland, Lithuania, Poland and Switzerland
From two groups in regular season	Georgia (top four in each group – 1999/00), Hungary (top six in each group – 2000/01), Poland (top four in each group), and San Marino (top three in each group)
Bosnia-Herzegovina	*1995/96 and 1996/97:* This championship playoff was a smaller type, a championship final between the winners of the two groups in the league (in the championship of Republic "Srpska", while the "Muslim" championship is not included in this championship playoff). *1997/98:* Championship playoff based on top four from the Muslim top division + top two in the Croat league (no teams from Republic of Srpska). Playoff: Two groups of three teams, where the winners went into a championship final. *1999/00:* Championship playoff based on top five from the Muslim top division + top three in the Herceg-Bosna league (no teams from Republic of Srpska). Playoff: Two groups of four teams, where the winners went into a championship final.

The qualification procedure in the NAML is much more complicated than in European football, since the system for qualifying for playoff is related to relative performance, both within the division and within the conference (league). Below, a short overview over the qualification procedures for the championship playoffs for the different NAML is presented.

Table 5.9: Qualification procedure for championship playoffs in the NAML

League:	Teams to qualify for championship playoffs:
NFL	Divisional winners + two best records in each conference = 12 teams
MLB	Divisional winners + best record in each league = 8 teams
NBA, NHL	Divisional winners + five best records in each conference = 16 teams

In Australian leagues (VFL/AFL and ARL NRL), the system is more similar to the European football leagues, when it comes to the structure of the regular season, since it is based on one league/division.

The third general difference among the leagues, when it comes to championship playoff systems, is the relationship between regular season and the playoff. This means the strength of the dependency between sporting performance in regular season and the starting point (seeding) in the following post-seasonal playoff tournament. Also here, great variation between leagues in European football can be observed. This relationship is related to both match scheduling and starting score. The following table shows the different relationships between the stages in European football leagues that have had championship playoff in the period 1995(/96) to 2004(/05).

Table 5.10: Relationship between regular season and starting point for championship playoffs in European football leagues

Relationship between regular season and playoff	League
Independent	Bosnia-Herzegovina
Only affecting the playoff schedule	San Marino
"Bonus points":	
- All points	Andorra (2001/02-2005/06), Armenia, Croatia (2000/01), F.Y.R. Macedonia, Finland, Hungary (2001/02-2003/04), Israel, Kazakhstan, Latvia, Lithuania, Luxemburg, Moldova, Northern Ireland, Scotland, Slovakia and Slovenia
- 50 % of the points	Bulgaria, Croatia (1997/98-1998/99), Estonia, Georgia, Malta, Poland and Switzerland
- Other amount of points	Andorra (2000/01 - 1/3 of the points)
- Bonus points on basis of standing of the regular season	Hungary (2000/01), Croatia (1995/96)
- Combinations	Serbia-Montenegro: A bonus point combination, where the standing in the regular season gives certain number of bonus points to the qualified teams, in addition to more bonus points on basis of the number of points received in the regular season
- Points (only) against the other teams that are qualified for the playoff	Azerbaijan

This overview of the relationship between regular season performance and championship playoff shows that in most of these leagues teams bring (all) their regular season points into the playoff. The second most popular method is to include 50 percent of the points. This type of relationship between the regular season and the playoff differs significantly from the NAML system, where this kind of relationship is much weaker. However, connections between regular season and post-seasonal playoff, beyond just qualification, can also be found in the NAML, where seeding procedures decide both which teams to meet in the playoff and home field advantage. Seeding systems are also a part of the Australian leagues.

Introducing a championship playoff, such as in the Scottish league, where the teams bring all points from regular season into the championship playoff, might not affect competitive balance at all.[183] Changes in incentive effects in the regular season should be limited, apart from possibly the teams around the last qualification place (number six out of 12 in the Scottish league). Neither should the distribution of championship winners over time be affected, because the only difference from a "normal" round-robin tournament is that the teams in the upper part of the table will meet each other one more time (equal for all the qualifying teams). Also for performance persistence, it is difficult to envisage any direct consequences. However, this might be the dimension that is most affected because of potential financial rewards from qualifying for the championship playoff, which over time might "create" a group of teams qualifying regularly, season after season, and hence increase the performance persistence (weaken the competitive balance).[184] In such a situation, greater difference between the better and the weaker half of teams might appear, and as a consequence, weaken win dispersion and performance persistence. On the other side, increased risk-taking behaviour as regards expenditure on playing talent among second and third tier teams in the division might increase the likelihood for more variations in performance across seasons. Calculations in Chapter 4 show that average performance persistence is lower in the last decade in the Scottish top division, compared to earlier periods.

[183] In other words, the championship winner wins on basis of the number of points received after 33 regular matches + 5 playoff matches = 38 matches in total.
[184] One of the main financial bonuses of qualifying for the championship playoff in the Scottish Premier League is another lucrative home match against one of the Old Firm teams (Celtic and Rangers).

Match Schedule

Another area of tournament structure that might affect competitive balance (and competitive intensity) is whether the match schedule is balanced or unbalanced. Pure round-robin (one home and one away match against all teams in the division) is the normal in regular seasons of European football. However, there are exceptions from this format. First, there are teams playing four times against each other (two matches home and away), and second, that some leagues operate with the teams playing three times against each other. For the season 2004(/05), the table below shows the European football leagues that are among these two groups.

Table 5.11: European football leagues where teams met each other more than twice in the 2004(/05) season

Number of regular season matches:	League:
3	Denmark (12 teams), F.Y.R. Macedonia (12 teams), Israel (12 teams), Scotland (12 teams)
4	Albania (10 teams), Estonia (8 teams), Georgia (10 teams), Latvia (8 teams), Lithuania (8 teams), Moldova (8 teams), Ireland (10 teams), Slovakia (10 teams), Switzerland (10 teams)

Related to balanced or unbalanced league structure, the only difficult point might be related to the home field advantage. In other words, for leagues where teams meet each other three times, a crucial question is if home field is decided randomly in the "third" part of the tournament, or if it is based on sporting performance. If the latter, it might be categorised as unbalanced schedule, and can be relevant in the context of competitive balance. Note that the leagues where the teams play each other four times (two matches both home and away) have a balanced match schedule.

The league structure in San Marino is unique in European football. For the 2004/05 season, the playing schedule was based on round-robin for a given team's group, plus one match against each of the teams from the other group. This means that one team from group A plays 12 matches within its group (of seven teams) and eight matches against teams from Group B (eight teams). The eventual imbalance here could have appeared if the home field advantage was related to sporting performance regarding matches against the teams from the other group.[185]

[185] However, in San Marino the teams do not have their own home fields. The fields are randomly chosen from a given number of stadiums (en.wikipedia.org/wiki/Campionato_Sammarinese_di_Calcio).

In the cross-sectional regression analysis below, the leagues that have had "pure" round robin all seasons in the analysis period are given the value one in the dummy variable, while the other leagues have this variable valued zero.

The system for playing schedule differs significantly between European football leagues and the NAML. In the context of competitive balance, playing schedule might be a relevant determinant, at least in the NFL, where 2/16 of the matches are based on results from the previous season. The match schedule for each team in this league is presented in the table below.

Table 5.12: Match schedule in the NFL over the seasons 2002 – 2009

One home and one away match against the other teams in its division	6 matches
All teams in another division in its conference, rotating in a three year cycle	4 matches
All teams from a division in the other conference in a four-year cycle	4 matches
Two intra conference matches based on previous season's performance[1]	2 matches
Total number of matches for each team in the NFL	16 matches

Source: www.nfl.com/schedules/tv/2006_opponents (which was later removed, therefore see en.wikipedia.org/wiki/National_Football_League for an up-to-date reference by April 2008).

[1] If first place last season, these two matches go against the two other first place teams from the two divisions that are not played against in its conference this season. The same for the second, third and fourth placed team.

For example the NHL has a different form of match schedule. For the 2005/06 season, this is shown in the table below.

Table 5.13: Match schedule in the NHL for the 2005/06 season

Own division: 5 teams meet each other eight times (home=away)		32 matches
Intra conf. – other div.: Against the ten teams four times (home=away)		40 matches
Inter conf.:	5 home games against teams from one designated division and five away games against teams from another designated division (annual rotation)	10 matches
Total number of matches for each team in the NHL		82 matches

Sources: www.cbc.ca/sports/indepth/cba/features/schedule_changes.html and en.wikipedia.org/wiki/National_Hockey_League.

The Norwegian ice hockey league had for the 2004/05 season "bonus matches" that came in addition to the "regular" balanced schedule, where all teams played four times against each other (36 matches).[186] The six bonus matches were split into three periods of two matches (one home and one away), which were decided on basis of the current

[186] Source: www.hockey.no (more specific; www.hockey.no/t2.asp?p=4916&x=1&a=105359).

standing prior to those matches. The two first bonus matches were played after round nine in the "regular" system. It was the "neighbour-teams" in the league standing that played one home and one away match against each other (number one against number two, number three against number four etc.). The same system was repeated after 18 regular matches (+ two bonus matches) and 27 regular matches (+ four bonus matches). One problem here might come from the "six point matches" as in the Norwegian ice hockey 2004/05, if one team wins all these matches, the system might actually worsen the (ex post) competitive balance.[187]

5.3.3. Result

Based on 50 of the 52 UEFA members (except Bosnia-Herzegovina and Lichtenstein), cross-sectional OLS regression analyses are done, analysing the model as is described earlier in this chapter. The results are presented in table 5.14.

The cross-sectional regression analysis is based on three main steps. First, all hypothesised determinants are included in the basic regression, except from a variable on the distribution of revenues from sale of broadcasting rights. The correlation matrix among the independent variables is presented in table 5.15. It shows that high correlation can be found between some of them, such as between population in a country and the number of citizens in its largest city (0.858). The correlation between the UEFA and the FIFA ranking is close to 0.7, and between the UEFA ranking and population of the largest city about 0.6. The size of the league has a correlation to the UEFA ranking of about 0.65. Among league structural variables on five season averages, both championship playoff and relegation, and the number of teams and pure round-robin tournament, have a correlation higher than 0.6. In other words, the general model presented for each of the three fundamental dimensions might have some tendency of multicolinearity. Therefore the second step is to follow a general-to-specific strategy (see, for example, Doornik and Hendry, 2006a), where variables with t-ratios less than unity are dropped, as long as the goodness of fit is not reduced and the diagnostics remain insignificant. Third, broadcasting rights are specially treated in post-regression analyses of residuals.

[187] There was a discussion about the bonus system in a Norwegian newspaper winter 2005. One claims that the reason for the introduction was that the league was too unbalanced, and that the bonus system should be a method to improve (ex post) competitive balance (www.adressa.no).

Looking first at the ten season average NSQF ratio (column I and II in table 5.14), four variables stand out as significant determinants. Initially, it is these four that have t-values above unity in the general regression (column I), and they are all significant in the second step procedure (column II) for explaining differences in within-seasonal competitive balance between the European football leagues. The sign of these variables (per capita GNP, the percentage of the agricultural sector, the FIFA ranking, and a dummy variable for largest city as a share of total population) are as expected. The proposed explanations of the key results from the ten season average NSQF ratio are discussed in the following.

Table 5.14: Cross-sectional regression analysis on determinants for differences in competitive balance across European football leagues

	NSQF ratio - 10 season		NSQF ratio - 5 season		SRCC – 10 season		SRCC – 5 season		HHI 10 s	1/HHI - 10 season	
	(I) All	(II)	(III) All	(IV)	(V) All	(VI)	(VII) All	(VIII)	(IX) All	(X) All	(XI)
Constant	2.1548^{3}	1.9305^{3}	2.0143^{3}	1.9781^{3}	0.6237^{3}	0.5198^{3}	0.7198^{3}	0.7764	0.5956	3.3951	2.4815^{3}
	(0.4748)	(0.1030)	(0.5003)	(0.4222)	(0.1760)	(0.0693)	(0.1820)	(0.1405)	(0.3535)	(2.1190)	(0.3819)
Population	-0.0006		0.0006		-0.0006	-0.0006	0.0012	0.0010	0.0023	-0.0302^{1}	-0.0305^{2}
	(0.0037)		(0.0046)		(0.0014)	(0.0006)	(0.0017)	(0.0016)	(0.0028)	(0.0166)	(0.0147)
Population density	0.0002		0.0000	-0.0274	0.0002^{3}	0.0003^{3}	0.0002^{1}	0.0002^{2}	0.0000	-0.0009	
	(0.0002)		(0.0003)	(0.0263)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0002)	(0.0010)	
Largest city	-0.0054		-0.0390		-0.0041		-0.0284	-0.0253	-0.0296	0.3074	0.3278^{2}
	(0.0429)		(0.0538)		(0.0159)		(0.0196)	(0.0181)	(0.0319)	(0.1912)	(0.1633)
Big share largest	0.1775	0.2145^{2}	0.1398	0.1493	-0.0168		0.0885	0.0819^{1}	-0.0158	0.0158	
	(0.1177)	(0.0803)	(0.1450)	(0.1343)	(0.0437)		(0.0527)	(0.0485)	(0.0877)	(0.5254)	
Small share largest	-0.0896		-0.1453	-0.1266	-0.0492	-0.0487	-0.0954	-0.0976^{1}	-0.0988	1.0839^{1}	1.0042^{1}
	(0.1310)		(0.1613)	(0.1201)	(0.0486)	(0.0352)	(0.0587)	(0.0560)	(0.0975)	(0.5843)	(0.5119)
Agriculture percent	-0.8802	-1.1428^{2}	-1.1594	-1.1932	-0.2298		-0.4839	-0.4180	0.4150	-4.4270	-3.7920^{1}
	(0.6848)	(0.4972)	(0.8072)	(0.7450)	(0.2539)		(0.2936)	(0.2563)	(0.5099)	(3.0550)	(1.9220)
GNP per capita	-0.0190^{3}	-0.0172^{3}	-0.0236^{3}	-0.0227^{3}	-0.0049^{2}	-0.0030^{2}	-0.0084^{3}	-0.0077^{3}	0.0001	-0.0027	
	(0.0059)	(0.0036)	(0.0068)	(0.0053)	(0.0022)	(0.0013)	(0.0025)	(0.0022)	(0.0044)	(0.0264)	
UEFA ranking	0.0027		0.0031		0.0012		0.0005		-0.0004	0.0097	
	(0.0039)		(0.0049)		(0.0015)		(0.0018)		(0.0029)	(0.0176)	
FIFA ranking	0.0033^{1}	0.0038^{3}	0.0049^{3}	0.0045^{3}	0.0005		0.0003		-0.0016	0.0123^{1}	0.0086^{2}
	(0.0015)	(0.0008)	(0.0018)	(0.0015)	(0.0005)		(0.0007)		(0.0011)	(0.0065)	(0.0041)
UEFA late	0.0107		-0.0070		-0.0093		-0.0562	-0.0497	0.0624	-0.3841	
	(0.1133)		(0.1377)		(0.0420)		(0.0501)	(0.0451)	(0.0844)	(0.5057)	
Championship PO	0.0204		-0.4013^{2}	-0.4111^{2}	-0.0960^{1}	-0.0670^{1}	-0.0966	-0.0666	-0.0395	0.1648	
	(0.1531)		(0.1688)	(0.1600)	(0.0568)	(0.0385)	(0.0614)	(0.0482)	(0.1140)	(0.6832)	
Share automatic rel	-0.3107		-1.2383	-1.3604	0.5861	0.5816^{1}	-0.0815		-0.4571	2.1922	
	(1.014)		(1.2560)	(1.1290)	(0.3761)	(0.3248)	(0.4568)		(0.7553)	(4.5260)	
Share autoPOrel	-0.0789		0.5536	0.6060	0.3048	0.3139	0.1708	0.0865	0.1300	-2.9873	
	(0.6648)		(0.6088)	(0.5629)	(0.2465)	(0.1912)	(0.2214)	(0.1797)	(0.4950)	(2.9660)	
Pure round-robin	0.0799		-0.2025	-0.2094	0.0036		-0.0456		0.0788	-0.4212	
	(0.1232)		(0.1522)	(0.1395)	(0.0457)		(0.0554)		(0.0918)	(0.5499)	
Average # teams	-0.0148		0.0238	0.0309	-0.0050		0.0079	0.0021	-0.0082	-0.0289	
	(0.0249)		(0.0316)	(0.0275)	(0.0092)		(0.0115)	(0.0075)	(0.0186)	(0.1111)	
Observations	50	50	50	50	50	50	50	50	50	50	50
R^2	0.60	0.57	0.55	0.55	0.41	0.37	0.42	0.40	0.15	0.25	0.17
F-test	3.407^{3}	14.76^{3}	2.821^{3}	4.193^{3}	1.56	3.571^{3}	1.638	2.319^{2}	0.40	0.747	1.858
Normality test	2.2086	2.6172	2.6156	2.5921	3.2814	3.7121	4.5222	3.5992	19.687^{3}	2.2304	0.8524
Hetero test	0.2700	0.5729	0.2752	0.5787	0.3906	0.5465	0.4856	0.7954	0.2260	0.1808	0.4532
Hetero-X test	N/A	0.4346	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	0.6866
RESET test	1.8829	2.4495	1.7750	2.3373	1.6931	0.6432	0.3557	0.3338	0.7697	0.6510	0.6510

Standard errors in parentheses; [1] Significant on ten percent level; [2] Significant on five percent level; [3] Significant on one percent level.

Table 5.15: Correlation matrix on the analysed determinants in table 5.14

	Population	Pop density	Largest city	Big share lrg	Small sh lrg	Agriculture %	GNP per capita	UEFA rank	FIFA ranking	UEFA late
Population	1.0000									
Pop density	-0.0473	1.0000								
Largest city	0.8580	-0.0582	1.0000							
Big share largest	-0.2767	-0.1704	-0.1795	1.0000						
Small share lrg	0.4693	-0.0223	0.1697	-0.2984	1.0000					
Agriculture %	-0.1083	-0.1952	-0.1056	0.1227	0.1486	1.0000				
GNP per capita	-0.0797	0.1772	-0.0668	-0.0107	-0.1562	-0.5800	1.0000			
UEFA ranking	0.5479	0.0806	0.6027	-0.2350	0.1811	-0.3248	0.2601	1.0000		
FIFA ranking	-0.4422	0.1423	-0.4889	0.2119	-0.2098	0.2598	-0.0909	-0.6957	1.0000	
UEFA late	-0.2910	-0.1495	-0.2942	0.2103	-0.1380	0.2862	-0.5183	-0.4420	0.4292	1.0000
Champ PO	-0.3973	0.0883	-0.4105	0.0000	-0.2414	-0.0311	-0.2470	-0.4799	0.5022	0.4899
Average # teams	0.5479	0.0520	0.5836	-0.4828	0.3302	-0.1082	-0.0026	0.6540	-0.5206	-0.3707
Share auto rel	0.2629	-0.2334	0.2869	0.1310	0.1928	0.1734	-0.2150	0.2285	-0.3136	-0.1589
Share autoPOrel	-0.2871	0.0098	-0.2988	0.0750	-0.1889	-0.0208	0.0565	-0.2010	0.2197	0.0561
Pure round-robin	0.4301	-0.0840	0.3706	-0.1162	0.3170	-0.1504	0.1876	0.4276	-0.2598	-0.3595
Champ PO_5	-0.2807	0.1837	-0.3000	-0.1288	-0.1971	-0.1706	0.0165	-0.2851	0.4152	0.1531
Av # teams_5	0.5442	0.0632	0.5757	-0.5067	0.2695	-0.1992	0.0919	0.6570	-0.5247	-0.4093
Share auto rel_5	0.2677	-0.3627	0.2976	0.0258	0.2039	0.2466	-0.2683	0.2358	-0.4315	-0.1022
Sh autoPOrel_5	-0.2486	0.0786	-0.2765	0.0388	-0.1641	-0.0570	0.1096	-0.1970	0.2862	0.0180
Pure RR_5	0.3768	0.1095	0.3154	-0.2103	0.2365	-0.1193	0.0572	0.3344	-0.2265	-0.2500

	Champ PO	Av # teams	Share auto rel	Share autoPOrel	Pure RR	Champ PO_5	Av # teams_5	Sh auto rel_5	Sh autoPOrel_5	Pure RR_5
Champ PO	1.0000									
Average # teams	-0.5230	1.0000								
Share auto rel	-0.4175	0.2847	1.0000							
Share autoPOrel	0.5414	-0.4633	-0.1826	1.0000						
Pure round-robin	-0.5604	0.5829	0.4017	-0.1897	1.0000					
Champ PO_5						1.0000				
Av # teams_5						-0.2285	1.0000			
Share auto rel_5						-0.5742	0.3596	1.0000		
Sh autoPOrel_5						0.6136	-0.3223	-0.4402	1.0000	
Pure RR_5						-0.2381	0.6150	0.2792	0.0077	1.0000

Higher per capita GNP improves competitive balance because of a bigger shared league market. A greater size of the agricultural sector might indicate that the population is more dispersed. The dummy variables related to the largest city as a share of the total population, give at least two interesting results. If this share is relatively low, it does not affect competitive balance significantly, while in the leagues where this variable is relatively high, it has a significant negative effect on competitive balance. A relatively big share of the total population in the largest city can therefore indicate that teams from these areas are more likely to dominate the league. The FIFA ranking can be interpreted as an indicator of the size of the football market, as is again likely to increase competitive balance. It is the only sporting variable of significance, when applying ten seasons average (column I and II). This means that neither the points achieved in the UEFA ranking system for the teams of a particular league, nor the age of the league,[188] can significantly explain the differences in win dispersion among the member leagues in the UEFA. The results suggest that variations in prize and tournament structure do not significantly affect competitive balance, measured by ten season averaged NSQF ratio.

The other variables are highly insignificant, and are hence not among the explanatory variables for explaining differences in ten season averaged within-seasonal competitive balance among the leagues in Europe. Compared to the descriptive results, where the Big Five leagues, on average, seem to be better balanced than other leagues, the insignificance of the total population variable might be surprising. Regression analysis, replacing the population variable with dummy variables for small and large total population leagues, does not change this conclusion. In other words, the total population for the different leagues in European football is not an important determinant for explaining differences in win dispersion, given the other variables in the regression analysis.

Reducing the number of seasons in the average NSQF ratio to five, which means the period from 1999(/00) to 2004(/05) (column III and IV), all population variables are insignificant. On the other hand, t-ratios are greater than one for the biggest city's share of the total population, both when this share is big and when it is small. Signs on these coefficients are consistent with hypothesis. The agriculture percent also follows this pattern, as insignificant, and the sign is still positive related to competitive balance.

[188] Using the initial membership year in the UEFA does not change this conclusion.

Different from the ten season period analysis. is that tournament structure seems to be more determining for the level of win dispersion. by t-values greater than unity. There is a tendency for leagues with higher number of teams to have weaker win dispersion, ceteris paribus, while a higher level of the share of teams to automatically relegate has the opposite tendency. Changing one team to directly relegate with a relegation playoff might have negative effects on win dispersion. Next, leagues having a pure round-robin match schedule have a tendency to have better within-season competitive balance than other leagues. Given the other variables, a league with championship playoff during the five seasons has significant better win dispersion than other leagues. The FIFA ranking and the GNP per capita are still strong significant determinants for differences in win dispersion across European football leagues, also in the reduced sample period.

Measuring competitive balance by the SRCC (column V to VIII) gives a less clear relationship between the explanatory factors and the dependent variable, as can, for example, be found by the weaker F-tests (insignificant when including the whole range of possible explanatory variables). Except from the GNP per capita, and partly the championship playoff variable, the other significant results from the regressions based on performance persistence on ten season average (column V and VI)) as the dependent variable, differ from the win dispersion regressions. For example, population density is a strong significant determinant for differences in performance persistence between European football leagues, where the persistence increases with higher density. The ten season average SRCC might be more (positively) affected by an increase in the size of the population (t-value slightly higher than unity in the second regression (column VI)). The indications that the (relative) number of teams being relegated, also when including the number of teams to relegation playoff places, have a negative impact on the comparison of performance persistence (ten season average) among the leagues in the UEFA area, are surprising in relation to the hypothesis about a positive relationship between relegation and competitive balance. Since none of the other regressions found any significant relationship between competitive balance and relegation, this might reflect that the importance for the differences between leagues is not much related to the relative number of teams into relegation and relegation playoff. However, it is impossible to interpret the effects on competitive balance from including relegation (open leagues vs. closed leagues). As mentioned before, since only one European football league is closed, it might be that this variable is more significant when analysing effects on competitive balance changes within leagues over time. This will be done in Chapter 6.

The results from the regressions based on SRCC averaged over five seasons (VII and VIII) confirm most of the findings of the ten seasons' average, except for the relegation variables. The results also show that variables related to the relative size of the biggest city are stronger determinants than in the ten season average performance persistence analysis. Higher relative size of the biggest city is (weak) significant increasing performance persistence, while small relative size of the biggest city is a (weak) driver for higher variation across seasons in sporting performance. Note that three of the variables in the reduced form have a t-value lower than unit. They are included because further reduction rejects diagnostics. However, a different specification with fewer variables, also without violating the diagnostics, gives much of the same conclusions (column VIII).

It is difficult to find a good model for the relationships between differences in championship concentration and the hypothesised determinants for differences in competitive balance across European football leagues (column IX to XI). This is confirmed by the diagnostics, for example showing much weaker R^2 and weak F-tests. Because of significant failure in the diagnostics when applying HHI "directly" as the dependent variable (column IX), it is also transformed into the inverse in the final equation. This means that the higher level of the dependent variable, the better dispersion of championship winners in column X and XI. However, even by these transformations (and others as well) the F-tests are still not significant.

The results from possible determinants for differences in championship concentration in column IX and X suggest that it increases by increased population and better FIFA ranking. These results are on odds with the hypotheses suggested for win dispersion. Even though the relative strong correlation between population and the size of the largest city, increased size of the latter variable increases the distribution of championship winners (significant in column X). Concentration is smaller in leagues where the biggest city is small compared to the total population (weak significant). Higher agriculture percent has a tendency to increase concentration, which is opposite from the results for win dispersion.

Residual Analysis in Relation to Differences in the Distribution of Revenues from the Sale of TV Rights among the Big Five Leagues

The tables below show the estimated residuals of the Big Five leagues with the countries ranked by the range in the distribution of broadcasting revenues. Italy is ranked first, as the league with the highest range, while France, as the league with the lowest range, is on the bottom.

Table 5.16: Residuals from the cross-sectional regression analysis based on win dispersion for the Big Five leagues

	NSQF10_all	NSQF10_sig	NSQF5_all	NSQF5_sig
Italy	0.1854	0.1480	0.2522	0.3217
Spain	-0.2289	-0.1424	-0.3982	-0.2917
Germany	-0.0246	-0.0792	0.0053	0.0677
England	0.0525	0.1123	0.1608	0.1887
France	-0.1186	-0.1517	-0.0273	-0.0723

The relationship between the residuals in Italy, Germany and France and the range of the distribution of revenues from sale of broadcasting rights fits relatively well, and might indicate that the distribution of revenues from sale of broadcasting rights could have been a relevant variable in the cross-sectional regression analyses. On the other hand, the residuals for the Spanish and the English leagues do not follow this pattern. This might indicate that there is no general relationship and/or that these leagues have something that is not captured in the same degree by the general regression model.

Table 5.17: Residuals from the cross-sectional regression analysis based on performance persistence for the Big Five leagues

	SRCC10_all	SRCC10_sig	SRCC5_all	SRCC5_sig
Italy	0.0533	0.0729	0.1208	0.1361
Spain	-0.0100	0.0208	-0.0641	-0.0367
Germany	-0.0881	-0.0748	-0.0117	0.0014
England	-0.0208	-0.0310	0.0257	0.0332
France	-0.0712	-0.1032	-0.0253	-0.0374

For the ten season average of the SRCC, the ranking of the residuals follows the ranking of the distribution of revenues from sale of broadcasting rights, except from the relatively better results for Germany. For the five season average the pattern is closer to the NSQF ratios above in the general function form, while the reduced form correlates between the distribution of broadcasting rights revenues and performance persistence, except from Spain.

Table 5.18: Residuals from the cross-sectional regression analysis based on championship winner concentration for the Big Five leagues

	HHI_all	1/HHI_all	1/HHI_sig
Italy	0.0223	-0.5221	-0.3478
Spain	-0.1830	1.4817	1.6909
Germany	-0.0122	0.1265	0.2522
England	0.1168	-0.9557	-1.1153
France	-0.0757	0.5873	0.4686

As for the win dispersion, the ranking between Italy, Germany and France correlates between competitive balance and distribution of revenues from sale of broadcasting rights, and Spain and England deviate from this pattern.

In summary, based on the Big Five leagues, there are some indications that differences in distribution of revenues from sale of broadcasting rights might be a significant variable in cross-sectional regression analyses of determinants for differences in competitive balance across European football leagues. For similar analyses in the future, it would be interesting to include this variable in the general analysis, as both the revenues to be distributed have become more significant in a higher number of leagues, and the distribution mechanisms have become more incorporated.

5.4. More about Competitive Balance and Championship Playoffs in European Football

A number of analyses focusing on possible relationships between competitive balance and championship playoffs in European football are presented below. The first part deals with comparisons of competitive balance in leagues without championship playoffs, compared to leagues with championship playoffs. The second part will concentrate on leagues that have had periods of both, and compare competitive balance in the different situations.

5.4.1. Competitive Balance and Changes in Championship Playoffs in European Football Leagues

The high number of European football leagues makes it possible to investigate relationships between prize/league structure and competitive balance to a higher degree than in many of the other main team sports in the World. As shown before, during the period from 1995(/96) to 2004(/05), just over 50 % (26 out of 51) of the European football leagues had the championship winner decided after championship playoffs in at least one season. An interesting research question is: what drives leagues to introduce championship playoffs? In the context of competitive balance, this is of high interest. Is weak competitive balance a driver for restructuring tournaments? On basis of the argument that competitive balance is one of the factors that affect changes in prize and league structures, the following hypothesis is made: Championship playoffs is, to a larger degree, used in leagues with weak competitive balance.

The hypothesis mentioned above is analysed by using the top divisions in 50 (51 for the HHI) of the 52 members of the UEFA. Following the structure of this thesis, competitive balance is calculated in three dimensions, using the measures applied earlier in this chapter. The competitive balance calculations for the win dispersion and performance persistence are made from the regular season data, and hence do not include post-seasonal playoffs. HHI is based on the championship winner (eventually after championship playoffs). The table below presents an overview of the calculated average values of the different dimensions of competitive balance.

Table 5.19: Comparisons of competitive balance in European football leagues with and without championship playoffs over the period 1996(/97) – 2005(/06)

	NSQF	SRCC	HHI[1]
Average CB "playoff leagues" (25 leagues)	1.961	0.580	0.360
Average CB "non playoff leagues" (25 leagues)	1.641	0.594	0.409
Difference	0.320***	- 0.014	- 0.049

*** Significant on one percent level (two-sided).

[1] The HHI is calculated on 26 playoff leagues (including Bosnia-Herzegovina).

The results from the t-tests are interesting, where the win dispersion seems to (significantly) differ between the leagues that are categorised as playoff leagues and the other European football leagues in the actual period of time. This result might strengthen the hypothesis that weaker competitive balanced leagues have an increased probability (for governing bodies) to introduce championship playoffs. However, since performance persistence and championship concentration are relatively equal in the two groups of leagues, this is only related to one dimension of competitive balance.

Further, analyses of the hypothesis that weaker win dispersion increases the probability for leagues to introduce championship playoffs can be done by regressions based on a dummy variable as the dependent variable. This binary variable is valued one if, during the period from 1995(/96) to 2004(/05), the league has had a period with championship playoffs (PO_i) and zero if not. Competitive balance (CB_i) is the independent variable. The following model is presented:

$$PO_i = \beta_1 + \beta_2 CB_i + \varepsilon_i$$

The results from binary logit analysis for the different measures of competitive balance are presented in the table below.

Table 5.20: Logit analysis of competitive balance as a playoff driver

	NSQF	SRCC	HHI
Constant	- 5.618***	0.744	0.680
	(1.986)	(1.636)	(0.708)
CB	3.135***	-1.267	-1.677
	(1.105)	(2.745)	(1.711)
Observations	50	50	51
Likelihood ratio test[1]	10.599***	0.214	1.023

Standard errors in parentheses. *** Significant on one percent level.
[1]Reference: Doornik and Hendry (2006b).

The logit model confirms the observations above, that the probability of having championship playoffs is highly significant in leagues with greater win dispersion.

The sign of the coefficient for performance persistence is negative (weaker performance persistence), but the coefficient has a t-value below 0.5, indicating that the probability for differences in performance persistence in leagues with championship playoffs, compared to leagues without, is highly insignificant.

For the dispersion of championship winners, the t-value close to one for the HHI might give some indication that there is a higher probability that a league with less concentration of championship winners is among the leagues that have had championship playoffs in the analysed period (opposite of the hypothesis). If the race for winning championships has a degree of sensitivity related to changes in tournament structure, this can be one possible explanation for the tendency in the results. The other is that the championship playoff in itself reduces the probability of creating dynasties for the champions. However, these hypotheses are not significantly confirmed in the calculations.

The interesting results from win dispersion need further analyses, because there might be other determinants for the relationship between competitive balance and championship playoffs among European football leagues. One obvious control variable is the age of leagues, which varies a lot, due to consequences of changes in (sporting) policy (inclusion of smaller leagues) and the general structural changes in the Eastern Europe. The argument for changes in league structure in newer leagues can be that time might be needed to find an appropriate structure. According to Sandy et al. (2004), instability is a general characteristic in new industries. The logit model above is therefore extended by the dummy variable UEFA, which is used for categorising the age of the league on basis of the membership year in the UEFA. New leagues are valued one, if membership in the UEFA is 1988 or later, and zero, if membership is earlier. This gives the following model:

$$PO_i = \beta_1 + \beta_2 CB_i + \beta_3 UEFA_i + \varepsilon_i$$

The empirical results from the binary logit analysis are presented in the table below.

Table 5.21: Logit analysis of competitive balance and age as playoff drivers

	NSQF	SRCC	HHI
Constant	-4.559**	0.082	0.023
	(2.172)	(1.892)	(0.816)
CB	2.197*	-1.592	-2.354
	(1.244)	(3.186)	(1.986)
UEFA	1.676**	2.245***	2.392***
	(0.740)	(0.690)	(0.712)
Observations	50	50	51
Likelihood ratio test	16.083***	12.900***	15.067***

Standard errors in parentheses. *** Significant on one percent level; ** Significant on five percent level; * Significant on ten percent level.

By including the control variable for age, win dispersion is now only a marginally significant determinant of a league having championship playoffs. Performance persistence is still not a statistically significant predictor for determining the probability that a league has had championship playoffs or not in the given period. The t-value of the HHI index increases slightly when including the dummy variable, representing the age of the league. In other words, an insignificant increased probability of being a championship playoff league can be expected in leagues with more dispersed concentration of championship winners. This is the opposite of the findings for win dispersion.

Compared to the calculations without controlling for age, it is the win dispersion that has the greatest difference between the models. When including age, it is only significant on a ten percent level, while it is on a one percent level without the control variable. However, the win dispersion level is still indicated to be a relevant variable for changing league structure into championship playoffs. When it comes to the age of the leagues, the hypothesis about more instability in the new leagues is highly significantly confirmed.

Even if the results from the logit analyses are confirmed in a probit model (not included in the thesis), they must be carefully interpreted. First, they do not say whether a season with playoff is better, worse or equally balanced, compared to seasons without playoffs. Second, the data do not divide between leagues, where one season was with playoffs, or if all seasons had this system. Third, the playoffs might in themselves affect competitive balance, for example through incentive effects, and by a changed structure. The latter can be a possible determinant, if championship playoff is introduced in leagues with less concentration of championship winners. In addition, the time period for the analyses is arbitrary.

5.4.2. Analysis of Championship Playoffs as Regular Season Competitive Balance Determinant

The results from the logit models above indicate that leagues with weaker win dispersion are more likely to include championship playoffs than other leagues. An obvious follow-up research question related to incentive effects on competitive balance, from the inclusion of championship playoffs, is: does the level of competitive balance differ in seasons with championship playoffs compared to seasons with no playoffs?

A difficulty in this kind of analysis in European football is that the relationships (incentive effects) between championship playoff and regular season competitive balance might be affected by the structure of the championship playoff. In some systems, one will expect regular season competitive balance to be relatively unaffected, while others might bring significant incentive effects, affecting regular season competitive balance. The incentive effects in regular season might be related to both if (parts of the) points won are transferred to the post-seasonal championship playoff, or if, eventually, seeding systems are related to regular season performance.

T-tests will be applied for five football leagues in Europe over a period of 35 seasons [1970(/71)-2004(/05)] to analyse if regular season competitive balance is affected by inclusion of championship playoffs. The football leagues in Sweden, Denmark, Luxembourg, Austria, Switzerland and Scotland all meet the following requirement: one top division in all seasons and significant period(s) with and without championship playoffs.

As shown earlier in this section, there are differences in the structure of the championship playoff systems in European football. The structure for the championship playoffs in the actual leagues are presented in the table below.

Table 5.22: Structure of championship playoffs in a number of European football leagues

League:	Playoff systems:
Sweden	1982-1984: Top eight (of 12 teams) to PO quarter finals. 1985-1990: Top four (of 12 teams) to PO semi finals. 1991-1992: Top six (of ten teams) to PO. Starting point: Half of the points from regular season. Six teams play ten matches.
Switzerland	1976/77-1978/79: Top six (of 12 teams) to PO. Starting point: Half of the points from regular season. Six teams play ten matches. 1979/80: Same as previous season, except that the total number of teams is 14. 1987/88-2002/03: Top eight (of 12 teams) to PO. Starting point: Half of the points from regular season. Eight teams play 14 matches.
Luxembourg	1987/88: Top four (of 12 teams) to PO. Starting point: Half of the points from regular season. Four teams play six matches. 1988/89-1993/94: Top six (of ten teams) to PO. Starting point: Half of the points from regular season. Six teams play ten matches. 1999/00-2004/05: Top four (of 12 teams) to PO. Starting point: All points from regular season. Four teams play six matches.
Denmark	1991/92-1994/95: Top eight (of ten teams) to PO. Starting point: Half of the points from regular season. Eight teams play 14 matches.
Austria	1985/86-1987/88: Top eight (of 12 teams) to PO. Starting point: All points from regular season. Eight teams play 14 matches. 1988/89-1992/93: Top eight (of 12 teams) to PO. Starting point: Half of the points from regular season. Eight teams play 14 matches.
Scotland	2000/01-2004/05: Top six (of 12 teams) to PO. Starting point: All points from regular season. Six teams play five matches.

Among these leagues, only the Swedish league has knock-out tournament championship playoffs (except in the 1990 and 1991 seasons). As shown earlier, the normal systems for championship playoffs seem to be a strong relationship between regular season point score and playoff league tournament, where either half or all points are brought into the playoff league tournament.

Below are shown average values of competitive balance during the analysis period of 35 seasons, dividing between periods with and without championship playoffs. Differences are tested by t-tests, based on the statistical null hypothesis and two-sided alternative hypothesis. The latter is chosen because incentives relevant for competitive balance are not clear in these leagues. The overview includes competitive balance calculations from the three dimensions used in this thesis. Note that the HHI is adjusted to take into account the number of seasons to be calculated. It can be called the relative HHI: rHHI = HHI/perfect HHI, where perfect HHI means HHI in a situation where no teams win more than one championship (perfect competitive balance), given the number of seasons to be calculated. Higher rHHI hence means a relatively higher relative concentration of championship winners.

Table 5.23: Comparisons of competitive balance when championship playoffs or not

League		NSQF	SRCC	rHHI
Sweden	Playoff	1.21	0.39	3.91
	No playoff	1.30	0.50	3.67
	t-values	1.135	1.478	
Switzerland	Playoff	1.31	0.48	4.50
	No playoff	1.76	0.63	3.40
	t-values	4.791***	1.976*	
Luxembourg	Playoff	1.72	0.62	2.69
	No playoff	1.64	0.57	9.64
	t-values	-0.786	-0.726	
Denmark[1]	Playoff	1.24	0.57	1.00
	No playoff	1.42	0.44	4.81
	t-values	0.987	-1.061	
Austria	Playoff	1.58	0.54	3.00
	No playoff	1.61	0.60	4.78
	t-values	0.182	0.764	
Scotland	Playoff	2.23	0.50	2.60
	No playoff	1.93	0.69	11.87
	t-values	-1.818*	2.330**	
Scotland[2]	Playoff	2.23	0.50	2.60
	No playoff	1.92	0.49	3.40
	t-values	-2.829***	-0.174	

*** Significant on one percent level; ** Significant on five percent level; * Significant on ten percent level. Two-sided t-tests.
[1] From 1971-2004/05.
[2] Alternative calculations over equal periods: 1995/96-1999/00 and 2000/01-2004/05.

Because of the low number of seasons with championship playoffs in the Scottish league, and that the value of the rHHI for the non-championship playoff period (because of strong domination) is very high, an alternative comparison period is also presented.

The results differ between the leagues. Competitive balance is significantly better when measured by the NSQF ratio and the SRCC in the periods of championship playoffs in Switzerland. In Scotland, different directions can be found in these two dimensions. There are different responses in the different leagues also for the rHHI. Luxemburg, Denmark, Austria and Scotland have all better rHHI when championship playoffs are used. One would also expect the Swedish league to be included in this group, since most of the championship playoff tournament was structured as knock-out competitions, but the values of rHHI is rather slightly the opposite.

The significant increased win dispersion together with lower performance persistence in the Scottish league, after including post-seasonal playoffs (also playoffs for the weakest half of the regular season), are interesting results. An intuitive explanation can be that the better teams are given an opportunity to increase the revenue stream, compared to the weaker teams, and hence, widen the dispersion of the quality among the teams. The increased level of variance in performance over seasons might be related to an increased level of financial risk-taking behaviour among the middle of the table teams, and hence, give increased variation. It might also be that some teams, who miss out on the championship playoff, but who are already safe from relegation (because all regular season points are transferred to the playoffs) lose motivation in the latter part of the regular season. On the other hand, calculations from the alternative comparison periods in the Scottish league might suggest that this is a result from a longer trend (towards less persistence), consistent with the results in Chapter 4. Further analyses of possible determinants for these results are seen as beyond the scope of this thesis. However, a number of factors relevant in time series of competitive balance will be discussed in Chapter 6. Note that these analyses might also be affected by incentive effects related to qualifications to the UEFA tournaments.

5.5. Appendix

Variables used in the analyses in Chapter 5 are presented in the two tables below:

Table 5.24: Competitive balance in three dimensions for the seasons 1995(/96) – 2004(/05)

	NSQF		SRCC		HHI
	10 s	5 s	10 s	5 s	10 s
Albania	1.579	1.969	0.611	0.604	0.540
Andorra	2.222	2.088	0.614	0.712	0.240
Armenia	2.547	2.605	0.619	0.630	0.260
Austria	1.750	1.562	0.639	0.595	0.200
Azerbaijan	2.142	2.253	0.601	0.565	0.360
Belarus	2.039	2.133	0.632	0.759	0.223
Belgium	1.745	1.877	0.594	0.632	0.300
Bosnia-Herzegovina					0.180
Bulgaria	2.033	2.173	0.737	0.830	0.240
Croatia	1.546	1.556	0.592	0.515	0.460
Cyprus	2.129	2.149	0.769	0.706	0.380
Czech Republic	1.496	1.477	0.589	0.546	0.520
Denmark	1.583	1.636	0.571	0.597	0.360
England	1.580	1.657	0.612	0.602	0.460
Estonia	2.381	2.846	0.733	0.852	0.380
F.Y.R. Macedonia	1.804	2.056	0.596	0.587	0.240
Faroe Islands	1.786	1.770	0.572	0.481	0.260
Finland	1.437	1.583	0.392	0.631	0.360
France	1.283	1.269	0.468	0.413	0.240
Georgia	2.030	2.051	0.434	0.651	0.460
Germany	1.412	1.464	0.476	0.568	0.420
Greece	2.031	1.969	0.710	0.793	0.680
Hungary	1.569	1.427	0.559	0.528	0.220
Iceland	1.247	1.100	0.514	0.404	0.300
Israel	1.781	1.697	0.608	0.610	0.260
Italy	1.631	1.683	0.669	0.649	0.360
Kazakhstan	2.248	2.200	0.506	0.547	0.260
Latvia	2.660	3.020	0.362	0.933	1.000
Lithuania	2.338	2.438	0.797	0.772	0.420
Luxembourg	1.828	1.770	0.617	0.491	0.420
Malta	2.136	1.919	0.832	0.799	0.340
Moldova	2.421	2.446	0.709	0.817	0.420
Netherlands	1.879	1.911	0.759	0.785	0.420
Northern Ireland	1.730	1.863	0.499	0.606	0.240
Norway	1.457	1.385	0.465	0.460	1.000
Poland	1.558	1.514	0.515	0.600	0.320
Portugal	1.654	1.625	0.648	0.584	0.420
Republic of Ireland	1.602	1.538	0.623	0.679	0.300
Romania	1.520	1.348	0.567	0.502	0.380
Russia	1.645	1.547	0.537	0.532	0.420
San Marino	1.775	1.875	0.446	0.540	0.220
Scotland	2.076	2.228	0.495	0.503	0.520
Serbia-Montenegro	2.023	1.886	0.604	0.671	0.460
Slovakia	1.618	1.409	0.506	0.239	0.220
Slovenia	1.566	1.374	0.694	0.668	0.580
Spain	1.381	1.341	0.605	0.567	0.240
Sweden	1.236	1.307	0.491	0.515	0.160
Switzerland	1.357	1.407	0.437	0.401	0.280
Turkey	1.784	1.743	0.665	0.584	0.420
Ukraine	1.890	1.929	0.505	0.522	0.680
Wales	2.009	1.984	0.549	0.614	0.540

Table 5.25: Basis for independent variables in the cross-sectional regression analysis

	Population	Pop. density	Pop Larg. city	GNP/ capita $	Agri- %	UEFA coeff.	UEFA member	FIFA ranking
Albania	3544808	122.79	270000	4500	0.475	1.831	1954	89
Andorra	69865	146.53	25000	19000	0.010	0	1996	147
Armenia	2991360	120.04	1322000	3900	0.234	2.165	1993	112
Austria	8174762	98.37	2072000	30000	0.035	23.375	1954	68
Azerbaijan	7868385	91.85	1700000	3400	0.141	1.165	1994	114
Belarus	10310520	50.10	1700000	6000	0.111	3.416	1993	90
Belgium	10348276	336.82	1122000	29000	0.019	28.5	1954	16
Bulgaria	7517973	74.13	1188000	7600	0.114	18.665	1954	34
Croatia	4496869	82.91	765200	10700	0.079	18.625	1992	20
Cyprus	775927	81.61	195300	19200	0.041	10.165	1962	98
Czech Republic	10246178	130.72	1233000	15700	0.031	27.95	1993	6
Denmark	5413392	126.36	1326000	31200	0.020	17.375	1954	13
England	49558800	380.40	6962319	28200	0.008	58.34	1954	8
Estonia	1341664	32.60	499000	12300	0.049	1.665	1992	67
F.Y.R. Maced.	2071210	81.37	448600	6700	0.113	3.497	1994	92
Faroe Islands	46662	29.35	13100	22000	0.270	1.165	1992	126
Finland	5214512	16.89	1163000	27300	0.043	7.208	1954	40
France	60424213	108.09	9600000	27500	0.027	43.468	1954	2
Georgia	4693892	72.69	1400000	2500	0.205	5.666	1992	93
Germany	82424609	234.86	3337000	27600	0.010	51.132	1954	12
Greece	10647529	81.86	3100000	19900	0.067	36.782	1954	30
Hungary	10032375	110.31	2017000	13900	0.033	12.79	1954	72
Iceland	293966	2.72	156000	30900	0.092	3.498	1954	56
Israel	6199008	282.82	2200000	19700	0.028	23.999	1992	50
Italy	58057477	192.96	4300000	26800	0.022	62.311	1954	10
Kazakhstan	15143704	6.30	1300000	7000	0.077	0.5	2002	135
Latvia	2306306	36.44	921000	10100	0.045	6.665	1992	53
Lithuania	3607899	54.98	553000	11200	0.061	3.998	1992	103
Luxembourg	462690	165.92	81800	55100	0.005	1.332	1954	153
Malta	396851	1192.51	99000	17700	0.030	2.998	1960	128
Moldova	4446455	133.67	765000	1800	0.210	5.832	1993	108
Netherlands	16318199	466.45	1150000	28600	0.025	33.498	1954	4
Northern Ireland	1696600	124.97	274000	23000	0.015	1.498	1954	124
Norway	4574560	14.42	507467	37700	0.025	19.575	1954	42
Poland	38626349	126.79	1609000	11000	0.031	21.625	1954	25
Portugal	10524145	107.86	1971000	18000	0.058	35.583	1954	17
Rep. of Ireland	3969558	52.74	993300	29800	0.050	3.331	1954	14
Romania	22355551	96.96	2130000	6900	0.131	12.957	1954	27
Russia	143782338	8.61	9299000	8900	0.052	21.041	1954	24
San Marino	28503	417.68	5000	34600	0.010	0	1988	162
Scotland	5054800	64.87	577869	27000	0.023	30.375	1954	51
Serbia-Mont.	10825900	103.06	1168454	2300	0.152	0	1993	41
Slovakia	5423567	110.58	448292	13300	0.059	9.332	1993	47
Slovenia	2011473	97.28	263290	18300	0.030	13.665	1992	31
Spain	40280780	78.43	4072000	22000	0.036	75.539	1954	3
Sweden	8986400	21.69	755305	26800	0.020	17.591	1954	19
Switzerland	7450867	182.94	984000	32800	0.015	26.25	1954	44
Turkey	68893918	85.11	8141163	6700	0.117	28.991	1962	8
Ukraine	47732079	82.51	2932000	5300	0.188	24.583	1992	61
Wales	2918700	140.78	309000	23100	0.015	2.165	1954	65

6. A Time-Series Analysis of Competitive Balance in English and Norwegian Football

6.1. Introduction

The descriptive analyses of competitive balance in Chapter 4 indicate changes in competitive balance over time in the English top division. Changes in competitive balance in the top division in Norway have been less significant, except for championship winner concentration. The pattern (for all three dimensions) for the Norwegian top division is more U-shaped, where competitive balance measures have better values in the middle of the sample period than both the first and the last periods. The main research question in this chapter is: what are the effects on competitive balance of changes in regulations and tournament/prize structure over time in two football leagues, the top divisions in England and Norway. These leagues from European football represent the currently most popular internationally (England) and a more average league (Norway).

The relationships between product and labour market interventions and competitive balance are among the most popular issues in the literature of the economics of professional team sports, and can be traced back to the invariance proposition in Rottenberg (1956). Since then, a number of studies have analysed these relationships, both theoretically and empirically. However, even if the theoretical analysis has been done both in North America and in Europe, there is a lack of empirical evidence from European team sports. Especially after the introduction of free agency, as a consequence of the Bosman verdict in 1995, empirical analyses should also be relevant for European team sports.

Changes in regulations and tournament/prize structure can have different sources. There are changes that are (only) related to the domestic governing body itself. This is for example typical for promotion and relegation systems. On the other side, a number of regulations are external to the domestic governing body, such as the number of teams to qualify for the UEFA Champions League, which is decided by an international governing body for the sport (UEFA). Labour market changes (e.g. the Bosman verdict) might be based on international regulations, while work permits are a matter for the domestic authorities. In other words, changes can be made both within and outside the sport, and both at domestic and international levels.

The time-series (regression) analyses in this chapter are mainly based on the win dispersion (NSQF ratio) and performance persistence (SRCC) dimensions of competitive balance. It is difficult using proper measures that reflect changes in the prize concentration, as a consequence of changes in regulation or structure of leagues in such analyses. The reason is that the measure for prize concentration uses up a high number of degrees of freedom (for example if using ten seasons as a period of measuring concentration, a lag of ten seasons would be required after each change in regression the time-series analysis), and is, therefore, omitted from the (time-series) regression analyses. However, the championship concentration (HHI) is applied in analyses based on averaging certain time periods and in analyses of moving averages.

6.2. Literature Review

Because of the lack of European evidence, the review of empirical studies is based on NAML. The consequences on competitive balance after the introduction of free agency in the MLB have been widely analysed. In general, the first generation of analyses were done by what Lee and Fort (2005, p. 158) call "short-term 'cross-section' type approaches". Here, competitive balance prior to changes in, for example, the labour market (such as free agency) is compared to the post-period competitive balance. This is among others done in Scully (1989), Quirk and Fort (1992), Fort and Quirk (1995) and Eckard (2001).

The next group of literature puts changes in labour market regulations, and also other structural changes, into time-series analysis. This is for example done in Maxcy (2002), Maxcy and Mondello (2006), Lee and Fort (2005), and Fort and Lee (2007).

More generally, the empirical literature analysing consequences of changes in regulations in the context of competitive balance can be split into three directions:

1. Pre- and post period t-tests (eventually simple regression models).
2. Time-series regression analysis by including relevant variables.
3. Structural break point analysis.

Most of the relevant literature from the first group comes from analyses of consequences from introducing free agency in the MLB. The overview by Szymanski (2003b) shows that results on competitive balance are mixed, but if there are effects, more studies find these to be rather positive (i.e. improved competitive balance) than negative (table 3, p. 1160).[189] Using the MLB as an example, Schmidt and Berri (2003) emphasise that empirical results of changed institutional rules might be dependent on the time period of analyses. These analyses from the MLB show positive or neutral effects on competitive balance after the introduction of free agency, and confirm the general results of improvement in competitive balance in the MLB since the early 1960s in Schmidt and Berri (2003), referring to studies, such as Quirk and Fort (1992), Butler (1995), Vrooman (1995), Horowitz (1997), Schmidt (2001), and Schmidt and Berri (2001). Eckard (2001) summarizes the findings from the previous literature in the following way (p. 433):

> ....empirical work to date has produced results that generally are mixed in direction and statistically insignificant. Overall, Rottenberg's (1956) invariance proposition is supported, that is, free agency appears to have had no significant effect on competitive balance.

In a time-series analysis of competitive balance in the MLB, Maxcy (2002) includes labour market dummy variables that might influence on competitive balance over time. Among these is a dummy variable for the period of free agency, and for periods reflecting "actions", such as owner collusion that might deviate from free agency behaviour. In addition, Maxcy uses both win dispersion and performance persistence as dependent variables, and the analysis is done over a longer time period (1951-1999). The effects on competitive balance of introducing free agency are in general either highly significantly positive (SRCC), or insignificantly positive (NSQF ratio). However, in the first period after the introduction of free agency (1976-1980, called the "reentry period" that is "capturing the effects of the reentry draft and implicit transaction costs" (p. 152)), the dummy variable outweigh this effect, indicating no change in competitive balance, compared to the pre free agency period for both methods.

[189] A number of these analyses include effects from rookie drafting on competitive balance. Since this chapter concerns football in Norway and England, discussions on this free market intervention are limited.

Schmidt and Berri (2003) claim that the improvements in competitive balance are not related to market intervention (consistent with the invariance proposition). This is supported by Lee and Fort (2005, p. 159), using break point detection techniques[190] for the sample period 1901-99 (based on two measures of competitive balance - the NSQF ratio and the excess tail percentages of the distribution of winning percents):

> Thus the draft, free agency, recent MLB expansion, and the growth in local TV revenue disparity do not coincide with shifts in competitive balance. Instead, we find statistically significant trends in improved competitive balance in each league over these time periods. This leads us to conclude that more gradual occurrences over time (more, and more geographically dispersed, population centres; diffusion of games through TV; and globalization of the talent pool) have played the dominant role in the behavior of competitive balance. .... The technique employed cannot tell us whether this improvement trend is because of or in spite of MLB efforts intended to enhance balance....

Schmidt and Berri (2003) focus on the improvement of the size of playing talent and the Gould hypothesis (p. 703) in their explanations of the positive trend in competitive balance in the MLB. They show that the Major League Baseball has had several significant expansions. The first was the inclusion of coloured players (according to Schmidt and Berri, 2003, p. 696, Jackie Robinson broke "the color line" in 1947), and the second was the increased level of foreign players. Moreover, there have been changes in the general population. Schmidt and Berri (2003) claim on p. 696 that: "At the beginning of the 20th century, when the people playing Major League Baseball were only white Northern American males, the population baseball could draw from was relatively small". Schmidt and Berri (2005, p. 415) conclude that:[191]

> The Gould hypothesis, therefore, argues that as the talent pool rises, greater player homogeneity should be observed.

Further, Schmidt and Berri (2003, p. 693) conclude:

---

[190] Analyses of structural break points on (English) football are done by Palacios-Huerta (2004), but not on competitive balance.
[191] Schmidt and Berri (2005) also show that the Gould hypothesis predicts a nonlinear relationship between population and competitive balance. See their page 415.

> Overall, our results suggest that the driving force behind Major League Baseball's improved competitive balance has been increases in the population of players Major League Baseball can employ. Although there exist marginal evidence of an impact from institutional changes, these vary and are quite small.

Hadley et al. (2005) concentrate their measure on competitive balance in the MLB on comparison of persistence related to playoffs pre and post the strike in 1994. After the strike, a number of changes took place, such as luxury tax, increased revenue sharing and expansions in the number of teams to post-seasonal playoffs. Despite these restrictions, they find competitive balance to be weaker in the after strike period. This is consistent with the findings about playoff concentration in Chapter 4 in this thesis. In addition, Chapter 4 also found the same significant pattern for AL (insignificant for the NL) in win dispersion and for MLB in championship winner concentration and performance persistence (both AL and NL).

Effects on competitive balance after changes in labour market regulations have received less attention in the empirical literature for the other major leagues in North America. However, a recent study by Maxcy and Mondello (2006) follows up the methods applied in Maxcy (2002) for these leagues.[192] The results are mixed, both with regards to league and measure. The estimated regressions cover the time period 1951-2004 in all leagues. In the NBA, competitive balance is insignificantly weakened by unrestricted free agency (based on both the NSQF ratio and the SRCC) compared to the initial period. The effects on competitive balance, as a consequence of salary caps, are dependent on measure. Both periods' competitive balances (1983-99 and 2000-04) are significantly weaker when using the SRCC compared to the initial period, while neither of these are significant for the NSQF ratio (however, the t-value for the first of the two salary caps periods is higher than one, negatively). The latter method confirms the findings in Fort and Quirk (1995).

In the NFL, the introduction of unrestricted free agency (1958-1963) has a significant negative influence on competitive balance in the study of Maxcy and Mondello (2006). The next period (1964-78, called the Rozelle Rule restricted free agency) is either weak

[192] A difference in the regressions in Maxcy and Mondello, compared to Maxcy (2002), is that Maxcy (2002) includes the effects of rookie drafting into the regression analyses.

significant or close to significant weaker than the initial period, when using the NSQF ratio, while when using SRCC, competitive balance goes back to the initial level. However, in one of the models, performance persistence has a weak significant improvement in the period from 1994 to the end of the sample period (2004), where payroll cap and unrestricted free agency were allowed. Larsen et al. (2006) also find that the combination of free agency and salary caps might improve competitive balance (weakly significant under the dHHI measure and insignificant under the standard deviation measure).

Maxcy and Mondello find the effects from introducing free agency in the NHL to (slightly) improve competitive balance when measuring competitive balance by the NSQF ratio. The restricted free agency period (1979-95) has a t-value higher than one and the unrestricted free agency in the period 1996-2004 exhibits a weakly significant improvement in competitive balance compared to the initial period. The performance persistence does not show any significant change.

The most recent analysis of the other major leagues has been undertaken by Fort and Lee (2007, p. 530), and they conclude that:

> While LF "Lee and Fort (2005) (sic.)" found a positive trend in balance in MLB, we find no trend in the NFL or NHL and a negative trend in the NBA. Employing break point techniques, we find no detectable structural change in within-season competitive balance prior to 1966 in these three NALs. This is in stark juxtaposition to the earlier findings for MLB where no break point occurred after 1962. So, we hypothesize that in one league (MLB), factors like the equalization of population centres, game diffusion on TV, and internationalization of the talent pool have been important in the determination of competitive balance. But in two leagues these factors have not been important (NFL and NHL). And in one league (the NBA), where the trend is negative, an additional explanation may involve the short supply of tall people, suggested by Berri et al. (2005).

These studies from the NAML show that there are different reactions in different leagues, when it comes to effect on competitive balance after changes in transfer market restrictions, as well as the response on different restrictions within single leagues.

Interesting in the context of this chapter, Maxcy and Mondello (2006) include the number of teams in the SRCC analyses. The effects on competitive balance are mixed across the leagues, but only the NBA shows a significant effect, indicating improvement in competitive balance. Also in the NHL the effects might be positive, but they are insignificant (t-value around one). For the NFL, the effects are rather opposite, with an insignificant decrease in competitive balance.

Schmidt (2001) analyses the effects on competitive balance in the MLB as a consequence of expansions, and concludes that (p. 21): "The results from estimating the time-series behaviour of league Gini coefficients indicate that the rise in competitive balance began with the movement toward expansion, i.e. 1962 for the American League and 1963 for the National League." However, these results might not be directly transferred to European football, because the procedures for expanding (or contracting) leagues are different.[193] In general, expansions and reductions of the number of teams in European football leagues are achieved through promotion and relegation based on sporting (not economic) criteria, namely, the previous season's rank.

Expansion has also been used in other empirical analyses of competitive balance after changing labour market restrictions, as a control variable for measuring effects in competitive balance. This is for example done in Fort and Quirk (1995), using dummy variables (only) for the seasons of expansions. A problem with regards to broader analyses of effects on competitive balance, after changing tournament structure, is that it is difficult to come up with general conclusions based only on the first season after expansion. Most studies find that competitive balance, measured as the distribution of sporting success, does not change significantly, but if there are any effects, they are usually negative. This is confirmed by the calculation of win percentages for the new teams in their first season in Schmidt (2001), with the average win percentage for an expansion team in its initial season being 0.371. Next season "these teams improve their winning percentage by nearly 10 %" (p. 21, footnote 1). When measuring competitive balance on basis of performance persistence, Butler (1995), Maxcy (2002), and Maxcy and Mondello (2006) find that if there are effects on competitive balance, they are positive. Larsen et al. (2006) find that the expansion year has no effect on NFL, while

[193] See Schmidt (2001) p. 21 for reasons why competitive balance might have been improved in Major League Baseball, as a consequence of expansion.

the number of matches affects competitive balance significantly (strong) negative, as the number of matches increases from 14 to 16.

Newson (1984), Dobson and Goddard (2001) and Fernandez-Cantelli and Meeden (2002) analyse effects of changing point score system, but these are not directly related to competitive balance.[194] It is therefore difficult to draw conclusions on competitive balance from them, because the aggregate distribution of home and away wins and draws are used. They do not consider the changes in the distribution of these results. Fernandez-Cantelli and Meeden (2002) claim there is a hypothesis about reduction in draws, when going from a (2,1,0) to a (3,1,0) system. This might be supported by Palacios-Huerta (2004), who finds a break point in draws as the point score system changed in England. However, empirically this is not generally supported as they (Fernandez-Cantelli and Meeden) find four out of ten leagues to have the opposite result. For a longer time period in England, a (slight) drop in draws might suggest that differences between teams have increased, but on the other side, the increased number of away wins improves the equality to home wins, which can be interpreted as positive in the context of competitive balance, if this means for all teams. However, if it means that the best teams now also win their away games, ceteris paribus, competitive balance is worsened due to the changed system. Comparing figure 1 in Palacios-Huerta (2004) and figures of competitive balance in the English top division later in this chapter, the frequency of draws and competitive balance might rather go in opposite directions, as is confirmed by calculating the correlation coefficient between the share of draws and the NSQF ratio (= 0.185) or the SRCC (= 0.336) for the whole sample period.

Related to Chapter 3, little research has been done on the effects on competitive balance from changed point score system in European football at team level. However, the game theory approach in Haugen (2008), which he also finds empirical support for in England, Norway and Romania, suggests worsened competitive balance (win dispersion) after changing the point score system. On the other hand, this is not consistent with Newson (1984, p. 91), who has the impression "that teams seem to move more rapidly up and down the table". This may be descriptive indications of

[194] For example, Dobson and Goddard (2001) find the number of away goals and away wins to increase in England, France, Germany and Spain in the period after introducing the (3,1,0) point score system, compared to the previous period. For the Scottish league, the number of wins was reduced. These results are not tested statistically.

decreased performance persistence, just after the introduction of the new point score system.

The NHL changed its score system prior to the 1999/00 season (see, for example, Abrevaya, 2004; Easton and Rockerbie, 2005). Before this change, the score system was (2,2,1,0,0), where wins counted regardless of regular time or in overtime. One point was (only) given if there was a draw in both regular time and overtime, while a loss gave zero points, regardless of whether after regular time play or overtime play. The new rule changed the number of points given, with a loss after overtime play awarded one point, i.e. (2,2,1,1,0) system". Both studies find that the changes in the overtime rules affect incentive effects.[195] This also happened, when an overtime rule was first introduced, prior to the 1983/84 season (Easton and Rockerbie, 2005).

When it comes to whether post-seasonal playoffs affect regular season competitive balance, Larsen et al. (2006) find different results regarding the way to define competitive balance in the NFL. By using dHHI as a measure of competitive balance, the effects are negative, as the number of teams to qualify increases.

6.3. Theoretical Framework and Rottenberg's Invariance Proposition

While Chapter 5 concentrates on determinants for differences in competitive balance across leagues, this chapter will focus on competitive balance effects from changes in regulations and prize structure within leagues over time. Therefore, only parts of the model in Chapter 5 can be applied in this chapter. Variables describing prize and tournament structure changes are relevant also in this chapter, while variables at macro level are omitted. The extension of the model in Chapter 5 is therefore mainly made in relation to a wider range of regulations, because of (within-league) changes over time. In other words, the framework for the analyses in this chapter is based on a split between regulations on one side and tournament/prize structure on the other. Regulations are further categorized into labour market and product market. Regulations

[195] Actually, another change also took place, since the overtime play is played by one player less on the ice for each team. However, transferring the knowledge from Abrevaya's (2004) analyses of the AHL (American Hockey League), where the similar changes took place during the last part of the 1990s, it was the changed point scores for the losing team in overtime play that was "responsible for the increase in the percentage of overtime games" (p. 298).

on the product market are gate revenue sharing and the distribution of revenues from collective TV-deals.

Three main labour market restrictions have been part of the history of European football. These are transfer market restrictions, limitations in the number of foreign players, and wage restrictions.[196] Among the arguments for these restrictions are that wages will be lower, and hence that smaller teams can compete for the better players, while player reservation rules will increase the smaller teams' ability to retain their best players, or at least ensure smaller teams to receive (higher) compensation for the players if eventually transferred (to bigger teams).

Historically, player transfers between teams have been strongly regulated, both in Europe and North America.[197] For example, player reservation systems[198] have "protected" teams against free movement of players at the end of their contracts, and in this way gave teams the right to retain their players and block their transfer to other teams or, if the team allows a transfer, to claim a transfer fee by way of compensation. In other words, in a player reservation system, a team still has a property right on the player's registration, even when the player is "out-of-contract". A free agency system, which is the current system in European football, implies that this property right only runs concurrently with the player's contract, and so that teams are, therefore, in general, unable to claim transfer fees for an "out-of-contract" player (apart from claiming compensation for training and development costs for younger players).

During the history of professional team sports, movement of players across leagues has been restricted. In Europe, until the Bosman verdict, the number of players from outside the country was generally strongly limited (e.g. three players).[199] One consequence from the Bosman verdict is free movement of players in professional team sports within the members and associated member countries of the European Union, and this labour

[196] Many of the same restrictions have been part of the NAML, additional to, for example, rookie drafting systems and luxury taxes. Descriptions of the latter can be found in for example Fort (2003) and Marburger (1997).

[197] Overview over transfer systems can be found in for example Quirk and Fort (1992), Szymanski and Kuypers (1999), Dobson and Gerrard (2000), and Dobson and Goddard (2001).

[198] Schmidt and Berri (2003, p. 694) claim that: ".....the reserve clause was enacted in a secret meeting at the National League in 1879 (Eckard, 2001)".

[199] Also the tournaments arranged by the UEFA have had restrictions on the number of foreign players. For the 2007/08 season this is related to "locally trained" players, "club-trained" players and "association-trained" players. See the Regulations of the UEFA Champions League, 2007/08 Article 17 (www.uefa.com/newsfiles/19071.pdf).

market can therefore be characterized as open. However, there are different practices when it comes to the different associations' behaviour to players outside Europe.[200] One can therefore say that European football leagues have gone from being a relatively closed to a much more open labour market during the 1990s.

Even if it is not directly involving revenue transfers, Downward and Dawson (2000) claim labour market policy can be treated as an indirect method to obtain cross-subsidisation, because it refers to the input market (the players). Further, they emphasise that this policy might not be that indirect, when looking at the sporting production process, since player expenses are a great share of a sporting team's costs (see, for example, Scully, 1989; Szymanski and Kuypers, 1999; annual reports from Deloitte). Changes in labour market policy can therefore affect the teams' finances at a significant level. However, since playing success is a function of the players' sporting capacity, labour market policy might affect the distribution of playing talent, and consequently, is a more direct method of affecting competitive balance.

Apart from governing body methods for interventions of free market solutions in professional team sports, competitive balance can be affected by owner objectives which may in some cases not correspond to profit maximisation. Késenne (1996) shows that there is weaker competitive balance under win maximisation with a zero-profit budget restriction compared to profit maximisation.[201]

Rottenberg's Invariance Proposition

The natural starting point in the theoretical analysis of competitive balance, from regulations on product and labour market, is the so-called *invariance proposition* from Rottenberg (1956). Szymanski (2003b, p. 1140) summarizes the main findings in Rottenberg:

> ....possibly the most important theoretical contribution to the analysis of team sports: the so-called invariance principle. This states that (a) changes in ownership rights over player services (such as the introduction of free agency)

[200] For example, the Norwegian Football Association increased the number of such players from two to three before the 2007 season after pressure from the teams.
[201] This could have been $\pi >= k$ as shown in for example Késenne (2006a), but the assumption of zero profit simplifies the analysis.

and (b) certain types of income redistribution (such as gate revenue sharing) will have no effect on competitive balance.

Given certain assumptions, Rottenberg concludes that the distribution of players is unchanged between a free agency system and a reserve clause system. In other words, Rottenberg claimed that the regulation would not have any effect on the distribution of playing talent.[202] Rich teams would purchase the best players, regardless of transfer systems, under the assumptions of a player transfer market (see, for example, Cairns et al., 1986; Eckard, 2001).

The literature finds the invariance proposition to be an example of the Coase theorem (Coase, 1960).[203] Noll (2006) claims that the invariance proposition is a special case of the Coase theorem (p. 28): "...that arises when transactions costs are unimportant. Unfortunately, the sports labor market is not an example of a market in which transaction costs can be ignored." Rottenberg uses some of the following assumptions:[204]

- Some teams have (much) larger revenues than others.
- Free labour market (the players accept the highest offers).
- Team objective is profit maximisation.
- At some point, the marginal return might be negative, which can happen if a team gets too dominant, and therefore reduces the uncertainty of outcome and, hence, the attendance.
- According to competitive balance, the richest teams in a league would, assuming that they are profit maximising, prefer winning by close margin. Therefore, regardless to regulations, playing talent would be relatively equally distributed among the teams.

[202] The owners claimed that given a free market for players, the big market teams would buy the best players and concentrate the distribution of playing talent among league members, which would result in lower uncertainty of outcome and reduction in league attendances (Cairns et al., 1986). According to Rottenberg, there is an important difference between a free agency system and a reserve clause system, since parts of the value of the service for the player is kept by the selling team (the transfer fee), while the player would get the whole value in a free market system. The literature is consistent about this wage effect for players.
[203] See, for example, La Croix and Kawaura (1999).
[204] Noll (2006, p. 21-22) discusses the relationship between Coase theorem and Rottenberg's invariance proposition, and claims that: "Rottenberg's analysis is not correct without further assumptions..." (p. 22).

The diminishing marginal return on additional team victories (higher win percent) within a season was introduced by Rottenberg. This effect was also a key feature of the first formal model on sports league, developed by El-Hodiri and Quirk (1971). Variants of models by El-Hodiri and Quirk (1971) have been used to analyse effects of different regulations in the sport market, among others the distribution of playing talent, and, hence, the competitive balance. The effects of market regulations have been discussed, both theoretically and empirically, by contributors such as Quirk and Fort (1992), Fort and Quirk (1995), Vrooman (1995, 1996, 1997, 2000), Hausman and Leonard (1997), Rascher (1997), Késenne (2000a, 2000b), Dobson and Goddard (2001), and Fort (2003). According to the summary of Dobson and Goddard (2001), regulations in the sports market, such as salary caps, maximum wage, reserve clause and gate revenue sharing, are not expected to give better competitive balance than the free market solution. On the other hand, these regulations may have implications for the distribution in revenues between stronger and weaker teams, and between teams and players.

Discussions of the Results from the Invariance Proposition[205]

The assumptions behind the invariance proposition have created a debate. According to Cairns et al. (1986, p. 31), Sloane (1976b) claims that a basic problem with the invariance proposition is "that the attendance-depressing effect of domination by a single team is largely external". This is related to the two variables, playing success and uncertainty of outcome, in a situation where one team is dominating the league. For this team, the two variables will go in opposite directions, while, for the other teams, they are going in the same (negative) direction. Cairns et al. (1986) summarise this effect:

> In a two-team league this externality could be internalised as described by Rottenberg, but in a many-team league this is less obvious. In terms of Rottenberg's example, teams C, D, E, F and so on may all enjoy benefits from the closer competition of B bidding for A's star player, but the market will fail to take these benefits into account.

Sloane (1969, 1971) argues against various assumptions in the proposition. Some of the arguments are summarised below:

[205] Restrictions on the number of players in a squad (Rottenberg, 1956) will not be treated in this thesis. However, this might be relevant in the future, because there are associations putting different restrictions on the squads when it comes to for example local players. For Norwegian football, see www.fotball.no.

- The profit maximising hypothesis among sports teams. Sloane claims that utility maximisation, given some financial constraints, might be a better assumption.
- Inter-league tournaments. In Europe, but also in the NAML, there are lucrative inter-league matches, where qualification is based on sporting quality.
- Only the rich teams are rich enough to sign the best players.
- Players do not want to leave a successful team for a less successful one.
- Sale of star players is unpopular among fans.

Even in a reserve clause system there will be unequal distribution of playing talent, especially when allowing for inter-team sales of players. According to Cairns et al. (1986, p. 32), Sloane (1976a, 1976b) argues that competition after all will be closer in a reserve clause system, since it will "slow down the rate at which star players gravitate to wealthier clubs".

On the other hand, Eckard (2001) shows that free agency might have a *positive* effect on the distribution of playing talent, compared to a reserve clause system. The argument is based on indications that the distribution of playing talent improved after introduction of free agency in MLB. The main reason for this hypothesis is that there "are diminishing marginal returns to each additional year's "production" of a championship calibre team" (p. 437). This is based on an argument related to an expected reduction in fan interest as the team continuously wins additional championships. Therefore, the team will have reduced incentives to continually bid for the best players, which is necessary for remaining at the top.

The influence of the objective of the owners/team in theoretical models

Sloane (1971) criticises the profit maximisation assumption in Rottenberg (1956) from a European football point of view. Traditionally, models from North America have mainly been related to profit maximising behaviour (El-Hodiri and Quirk, 1971; Quirk and Fort, 1995), while models from European team sports have questioned this assumption and have also analysed alternative ownership purposes into the models, such as in Késenne (various). In general, Noll (1982) shows that if teams' objective is to maximise the number of matches won, the player reservation system "prevents the best-financed team from acquiring the players whose contracts are the exclusive property of other teams, at least unless it obtains their acquiescence" (p. 387). On the other hand, Késenne (1996) supports the invariance proposition, even for leagues where the owners are maximising wins, subject to break-even. However, Fort (2006b) claims that in this case, the "invariance principle need not hold" (p. 85). Compared to the profit maximising models, Késenne (1996, 1999, 2006a) shows that competitive balance in a win maximising model is weaker.

One phenomenon is that "special" owners contribute to an abnormal positive cash flow, through supporting the given team to achieve sporting objectives, rather than financial payback. A recent example is the so-called "Abramovich effect" for Chelsea FC in the English Premier League.

Cost restrictions (salary/payroll caps, maximum wages)[206]

A problem, when it comes to analyses of competitive balance effects from salary caps, is the number of variants in practice (Késenne, 2002a). Theoretically, effects on competitive balance from introducing maximum wages are mixed.

Fort and Quirk (1995) argue that, theoretically, perfect competitive balance may be achieved through salary caps in combination with the reverse-order-of-finish draft system. However, one problem is that there might be incentives for teams to move away from this situation. This is because, at this equilibrium point, the marginal value of talent for the large market team is higher than for the small market team in a hypothetical two-team league (Dobson and Goddard, 2001).[207] Such a situation will motivate the large market team to buy player(s) from the small market team, and thereby diverge from the assumed effect of the salary cap restriction. Fort and Quirk (1995) claim that this is one explanation of the "ineffectiveness" of the salary cap in NBA, which is, according to Fort (2000), supported by Staudohar (1999), who finds that, on average, most teams have been above the cap in the 1990s. Vrooman (1995),

---

[206] Introduction of a salary cap system might include administrative problems, such as monitoring costs and enforcement problems (see, for example, Quirk and Fort, 1995; Fort, 2000). Rottenberg (1956) claims that a salary cap is not effective if there are possibilities for players to receive non-monetary goods, or that there are other ways to evade rules. Even if the system is working as anticipated, Rottenberg still claims that the players want to join the wealthiest teams.

[207] Dobson and Goddard emphasise that this solution is also inefficient in relation to maximising total league revenues (see also Quirk and Fort, 1992). This effect can be even stronger if the result is reduced TV rating, as a consequence of relatively weaker performance among the strong drawing power teams (Fort and Quirk, 1995).

among others, claims that because of zero marginal costs for talent (the costs do not go up with winning percentages), the profit maximising solution will be at the point where the league maximises its revenues. Interpreting this view, the same equilibrium point as free agency competitive balance is expected, and, as a result, salary caps should have no effect on competitive balance. On the other hand, Késenne (2000a) shows that a salary (payroll) cap improves competitive balance.[208] Even if a consequence is that the market move away from Pareto optimum, Késenne argues that it can be justified, because the market (and the owners of the bigger teams) may not take into account the reductions in negative externalities by improving competitive balance. Further, Késenne (2000a, 2002a, 2007b) finds that including individual cap most likely brings competitive balance closer to the free market solution, which means weaker competitive balance than the "pure" payroll cap solution.

Open labour market

Assume that a two-team league has a total stock of talent equal to $T_1 + T_2$. Given a closed labour market, the sum will be equal to unit, so if one team increases its stock of playing talent, the other must, by definition, decrease its number. This might be different in an open labour market regime, where one team can increase its stock of playing talent without affecting the other team (buying a player from outside the league), or that the other team can replace its loss of playing talent (after a within league transfer), by getting players from outside the league. These two examples give $T_1 + T_2 > 1$ after the changes. Therefore, compared to an open labour market, a closed labour market will, ceteris paribus, distribute playing talent more unequally. The reason is that acquisition of one extra talent does not necessary weaken the other team in an open market, so the effects on win ratio and revenue for the acquiring team will be less than in a league with a closed labour market. Hence, one can expect lower domination in an open labour market (Dobson and Goddard, 2001). Therefore, improved competitive balance is to be expected in domestic leagues in the post-Bosman period (more open labour market), compared to the pre-Bosman period (more closed labour market), ceteris paribus. However, this is not supported by theoretical models in Késenne (2007a), claiming no changes in competitive balance within leagues after opening the labour market. The market for players should also be bigger in an open market, as might also extend eventual competitive balance effects, as described in the Gould hypothesis.

[208] According to Késenne (2007b) this also holds when assuming win maximisation.

What is important in this discussion is the structure of a league, and at which level the labour market is closed. In league systems organised as divisional merit hierarchy, such as in European football, a top division team may buy a player from a lower division team, and, hence, not affect the playing squad for other teams in its own division. A horizontal structure of a league has some of the same effects, such as in the NAML, but it might be dependent on match schedule. For example, transferring a player from a team in another conference will not affect the playing squad for the teams in your own division. However, it can be of relevance in eventual playoffs.

Gate revenue sharing policy

If revenue sharing policy is successful in the context of promoting competitive balance, the revenue dispersion among the teams should be relatively better than the distribution of drawing power. According to Dobson and Goddard (2001, p. 130), revenue sharing policy in the NAML has been important with regards to "offsetting the basic inequalities in drawing power between teams". In Europe, revenue sharing seems to be less common, especially when it comes to gate sharing.[209] One reason might be weaker incentives for revenue sharing in open leagues (Szymanski, 2003b). However, even if European football leagues are open, incentives might be more driven by competitiveness in lucrative international UEFA tournaments than the relegation system.

For example Quirk and Fort (1992), Fort and Quirk (1995), Vrooman (1995), Dobson and Goddard (2001), and Késenne (2006a) analyse theoretically how revenue sharing affects competitive balance in a league with profit maximising owners. When assuming concave revenue functions (the analyses are based on a unitary league system,[210] the same proportion of revenue sharing, and ignoring local TV revenues), competitive balance will be unaffected from gate revenue sharing. This is robust for an n-league model (when ignoring local-TV revenues) (Fort and Quirk, 1995).

However, theoretical effects of gate revenue sharing policy, with regards to the distribution of playing talent, are dependent on the assumptions of the models (Késenne,

[209] See for example Table 1 in Buzzacchi et al. (2003).
[210] No playoffs.

2004, 2005a, 2006a). Based on the same assumptions as in the models above, except from profit maximising behaviour, Késenne (various) finds that revenue sharing improves the distribution of playing talent in a league consisting of win maximising teams. Moreover, Késenne (2006a) claims that competitive balance is improved regardless of specifications when win maximising behaviour, as is also shown in his many contributions.[211] Szymanski and Késenne (2004) relax another assumption, the fixed supply of talent, and find revenue sharing to weaken competitive balance in a two-team model. In the discussion of the general model, Noll (2006) mentions that the concave revenue function can be "disturbed", because leagues create "winners-take-all" systems.[212] Noll uses England as an example, where the top teams can earn substantial revenues, by qualifying for pan-European team tournaments. According to Szymanski (2001), Szymanski (1998) argues that by changing the shape of the revenue functions to be convex (by increased win percent), weaker competitive balance can be expected, given certain assumptions. However, Sloane (2002) questions the realism in incentives related to the model by Szymanski.

If all revenues in a league are transferred into a pool, which next distributes them equally among the teams, one would expect equalisation of playing talent among these teams, given a transfer market and profit maximization (Rottenberg, 1956, see also Cairns et al., 1986). However, this situation gives no incentives to win, and each team will try to minimise costs (Rottenberg, 1956). According to Cairns et al. (1986), Sloane (1980) claims the joint revenue will decline, because this will reduce attendance and the interest for the league. Sloane further claims that the resistance among the richest teams would also be a problem with such a policy. Késenne (2005a) analyses effects where parts of each team's revenues go into a pool. The n-team pool-sharing system might have negative effects on competitive balance when profit maximising objectives, while it is positive when win maximisation.

General comments

Dobson and Goddard (2001) summarize the theoretical implications of regulations in the sports market.[213] The main conclusion is that sports economists, in general, claim

---

[211] See also Késenne (2006b).
[212] Noll uses Frank and Cook (1996) as reference on the "winners-take-all" systems.
[213] The theoretical literature concerning competitive balance is mainly developed by US sports economists, typically with NAML as models (Dobson and Goddard, 2001).

that a free market regime in a professional league will "maintain a reasonable degree of competitive balance" (p. 125). This is argued by the interdependence between the teams in a league, which should motivate the "richest" teams not to concentrate the playing talent too much, so that the league does not lose its "competitive viability" (p. 125). In other words, the teams in a professional sports league have an economic incentive that the league should not be too unbalanced, or that superior playing talents are concentrated in one or a few teams (see also Rottenberg, 1956). On the other hand, every team also has a financial motive to be better than the other teams in the league, according to El-Hodiri and Quirk (1971, p. 1306), because "gate receipts of the home team are an increasing function for the probability of the home team winning for some range beyond a probability of .5". Moreover, positive financial payback from qualifications to post-seasonal tournaments might increase these motives. Szymanski (2003a, p. 471) shows that:

> The long-term nature of fandom in team sports has made the question of competitive balance central to policy analysis. Organizers of sports leagues, especially in the USA, have argued that to preserve the long-term interest in the game it is necessary to redistribute resources in such as a way as to ensure competitive balance.

While most of the methods for promoting competitive balance are questioned, with regards to reaching their purpose, the regulations from governing bodies have, in general, positive financial effects for teams and owners. Therefore, researchers question what the governing bodies' real incentives are for introducing market regulations.[214] Are the competitive balance arguments given only as an excuse to cover their financial incentives for regulations? However, there is also an unanswered question why the NFL is the best balanced NAML, if regulations do not have any effects on competitive balance. Among the NAML, the NFL is the most regulated (see, for example, Sandy et al., 2004).

Other theoretical effects on competitive balance, related to variables applied in this chapter, are presented in Chapter 5. In this chapter, Rottenberg's invariance proposition will be extended into changes in both regulations and structures. In this way, the

[214] This is emphasised in Balfour and Porter (1991, p. 10), who claim that "A major reason to question the competitive balance arguments comes from the fact that it is highly self-serving from the owners".

statistical null hypothesis will be applied for the analysed changes during the history of English and Norwegian football. The discussions will therefore only be related to deciding whether the alternative hypotheses are one- or two-sided.

6.4. Changes in Regulations and Tournament/Prize Structure in the Context of Competitive Balance. The Case of the English Top Division

The English football league tournament was established in 1888 (the 1888/89 season), when 12 teams played a round-robin tournament (one game home and one game away against all other teams in the division). A win was valued by two points, a draw with one, while losing a game did not give any points at all. Preston North End became the first championship winner. Automatic relegation to a lower division was not yet introduced. However, the poorly performing teams were not assured of retaining their membership of the English football league, since the bottom four had to apply for re-election, even before the second division was implemented from the 1892/93 season.

For a long period of time, the top division only had two prizes; winning the championship and avoiding relegation. In other words, knowing the many prizes that can be won by sporting performance in the current system of the English FA Premier League, the competitive intensity, ceteris paribus, was much lower at the start of the sample period in this analysis.

The English league is well documented when it comes to changes, both in structural and regulatory variables, as will be presented in the following.

6.4.1. Regulatory Changes in the English Top Division

6.4.1.1. Changes in Labour Market Regulations

The relationship between labour market policy and competitive balance has been central in the history of professional team sports (see, for example, Rottenberg, 1956). It has been related to restrictions in the transfer market, on wage payment and on limitations on the number of foreign players in English football. Labour market restrictions are also a part of the early history of the English football league, as shown in Green (1953, p. 405):

> In fact, what The Football League set out to achieve as soon as humanly possible was first to fix a maximum wage, and then the offer of a reasonable rate of payment by a club sufficient to entitle it to place a player's name on the retain list. In this way it was hoped that the greedy clubs would be taught that nothing was to be gained by underground methods of inducement, and that if they wanted a player they must be prepared to pay a transfer fee sufficient to enable the denuded club of replenishing this strength.

The labour market policy in the history of English football is interesting in the context of competitive balance, both because of a long period with maximum wages, but also because of a number of changes in the transfer system.

Transfer Rules[215]

Dobson and Gerrard (1999, p. 260) show that players in English football have to be "registered with the sport's governing bodies", to be given permission to play. Only one team can hold a player's registration at a certain point of time. If a player moves to another team, the player registration has to be transferred from the old to the new team. Player registrations are valuable for teams, and, for example, a market for transferring registrations (and hence players) appeared already in the early days of English football. Today, the transfer market is a natural part of professional football. For example, the 14 top teams in the 2000 season in Norwegian football transferred 95 players into their teams from other teams (Ohr and Solum, 2001). According to Dobson and Gerrard (1999), after the introduction of the freedom of contract transfer system in 1978, per season average permanent transfers were 294, which is about 15 percent of the total number of registered full-time professionals in English football.[216]

[215] The expressions used in the brackets below might not be the 'official' terminology applied, but will be the terms used in this thesis.

[216] For more information about the number of transfers in English football, see table 1 in Dobson and Gerrard.

The transfer market regimes have changed a number of times in English football. In 1885, three years before the foundation of the English Football League (FL),[217] the English Football Association (FA) accepted professional football (Green, 1953).[218] As a part of this acceptance of professional football, the relationships between the employer (football team) and the employee (the player) were regulated, as a registration system for players (Green, 1953). From the Football Association Constitution of 1887, Green (1953) shows that paragraph 26 stated that all players had to be "annually registered on a form to be supplied by the Secretary of the Association" (p. 182). According to paragraph 25, transfers during the season were restricted; "No professional shall be allowed to play for more than one Club in any season without special permission of the Council of the Association" (p. 182 in Green, 1953), which meant that transfers could not, in general, be given during the season, but even more important for an academic discussion of this transfer system in the context of competitive balance, is that, as shown in Green (1953) in a discussion from 1899, players were free after the end of each season to "enter into an agreement with any club for the following season" (p. 407). In other words, in the terminology applied in the current system of professional team sports, the initial transfer system in the English football league was a free agency system, which will be called "the FA free agency" in this chapter.

This "transfer system", created by the Football Association, still remained during the first seasons of the Football League. It was not until the 1893/94 season that the Football League introduced its "own" registration system, the so-called "retain and transfer" system,[219] as an answer to the failure of getting agreements on wage restrictions (Green, 1953). This registration system was much more restrictive than the FA registration system.[220] Given certain conditions, a player was not free at the end of contract to choose an employer. In other words, a tight (one-sided) binding agreement between the team and the player limited a player's possibility to play for another team, as long as the current team wanted the player. A player was not allowed to play for other teams, if the former team did not release the "player registration". This transfer

[217] For more about the foundation of the English Football League in 1888, originally a union of 12 teams, see, for example, Green (1953).
[218] The English Football Association was established in 1863 (Green, 1953).
[219] The English Football Association was not happy with this system (Green, 1953). Green (1953, p. 406) emphasises that: "The fundamental difference between the outlook of the F.A. and the League in the early days was that while the former put the interests of the individual and his freedom first, the League set out from the beginning primarily to protect the club. Therefore it became a question of the 'Individual' versus 'the System'."
[220] As can be read from the High Court case of Eastham v. Newcastle United Football Club Ltd. and Others [1963], players needed to be registered both at the FA and the FL.

system was, according to Szymanski and Zimbalist (2005), very similar to the reserve clause that existed in the MLB at this time. It lasted until the High Court case[221] involving the Newcastle United player Mr George Eastham in 1963.[222]

Eastham won the case against his former team, and the previous system was seen to "be an unreasonable restraint on trade" (Dobson and Goddard, 2001, p. 92). The new system, which can be called the "option-and-transfer" system,[223] was a modification of the former. New contracts had to be at least as good as the previous, both with regards to time (one or two years) and financially, if the team was to claim the rights to keep the player registration after the contract had expired (Sloane, 1969; Dobson and Goddard, 2001). In other words, as described in Sloane (1969), the structure of the transfer market regime did not change much in the new system.[224] Further, McArdle (2000, p. 13) emphasises that:

> If the club was unwilling to do that, the original contract would continue to run until he was transferred.

According to McArdle (2000), an independent Transfer Tribunal for arbitration was another consequence, if no agreement between team and player was made. McArdle therefore claims that (p. 13): "The days of clubs being able to bring players' careers to an end if they refused to accept whatever terms the club offered were over".

The "option and transfer" system lasted until a new system named "freedom of contract" was effective from the 1978/79 season. It was, according to Szymanski and Kuypers (1999), introduced in April 1978.[225] "Freedom of contract" was a further step towards free agency, even though it cannot be categorised as a free agency system. On one side, the freedom was established for the players, since they were now able to decide which team to play for, after contracts had expired. However, on the other hand,

[221] Mr Justice Wilberforce J.'s interpretation of the "retain and transfer" system can be found in the case Eastham v. Newcastle United Football Club Ltd and others [1963] 3 W.L.R., p. 583-4.
[222] Nawrat and Hutchings (1998) write the following for the 1960-61 season (p. 108): " ....Ministry of Labour on January 18, the League finally agreed to abandon the regulations that effectively tied a player to one club for life. The players' victory was finally confirmed when Eastham's case reached the High Court...".
[223] From lecture notes by Gerrard (2000b) for the course "The economics and finance of Football" at the Leeds University Business School.
[224] See Sloane (1969) for more detailed descriptions.
[225] Nawrat and Hutchings (1998) p. 214 show that: "When the details finally emerged in April,.......John Lacy became of the first players to put the new process to the test when he moved from Fulham to Tottenham in July". This is from the 1977-78 season.

free agency was not "achieved" yet, because the freedom was related to the behaviour of the former team. If the former team offered a contract at least as good as the previous, the team could claim a transfer fee. This transfer fee could be decided after negotiation with the new team, or if no such agreement, the "case" would be sent to an "FA tribunal". If the former team was not able to offer the kind of contract described above, the player was free to move without transfer fees (on free transfer) (Dobson and Goddard, 1998, 2001).

This system lasted until the Bosman verdict was implemented in English football, which happened in two phases. First, from 1996 transfer fees for an out-of-contract player moving to a new team, were seen to be of "divergence" with "Article 48 of the Treaty of Rome for the freedom of movement of labour" (Dobson and Goddard, 2001, page 95). The direct result was that out-of-contract players moving across borders within EU/EU-associated countries could do so, without the former team claiming transfer fee, and that strict limitations of foreign players was cancelled for players from EU (included associated members). This was a significant change, also in English football, where the limitation had been three foreign players (Magee and Sugden, 2002).[226] The second phase is related to the introduction of the domestic part of the Bosman verdict in July 1998 in English football. This stated that players of 24 years or older should now be free agents when the contract had expired (Morrow, 1999).

Magee and Sugden (2002, p. 421-422) analyse international migration in professional football, with a focus on England.[227] Their explanations of the increased number of foreign players by the millennium change are, apart from the Bosman verdict, the significant improvement in revenues through the UEFA Champions League and media and sponsors. Based on different sources, trends in the share of foreign players in the different Big Five leagues can be presented:

[226] This was due to regulations from the UFEA (3+2 rule). See, for example, Pearson, G. (University of Liverpool FIG Factsheet) - www.liv.ac.uk/footballindustry/bosman.html, accessed by April 2008.
[227] Magee and Sugden (2002) give an overview over typology for determinants for migration of players. They refer this terminology to Maguire (1996, 1999). Further analysis is beyond the scope of this thesis.

Table 6.1: The share of foreign players in the Big Five European football leagues in the 1999/00 and 2004/05 seasons

League	1999/00[1]	2004/05[3]
English FA Premier League	45 % (1998/99)[2]	56 %
Italian Serie A	33 %	31 %
German Bundesliga	40 %	48 %
French Ligue 1	23 %	38 %
Spanish Primera Liga	40 %	29 %

[1] Source is according to Magee and Sugden (2002): Massarella (2000).
[2] Source is according to Magee and Sugden (2002): Morgan (1999).
[3] Source: Gratton and Solberg (2007). They refer to football-europe.com.

In other words, opening the borders between the different nations within the European Union, when it comes to the labour market of professional team sports, led to expansions in the "population of athletes",[228] at least for the big revenue leagues, in the context of competitive balance. This is of interest in relation to the Gould hypothesis (Schmidt and Berri, 2003), described earlier in this thesis.

The current transfer system (Bosman II) was introduced prior to the 2002/03 season in English football, and is based on an agreement between FIFA, UEFA and the EU. It is an adjustment to the "Bosman I", by formalising fees for younger players (under the age of 24), and a number of other restrictions for these players. Some formal statements about length of contract and age of players are agreed on as well (Gerrard, 2002). The greatest change is probably the introduction of the so called "transfer windows", which is the general time periods when players can be traded between teams (the immediate consequence was that almost all changes between teams happen during these time periods, and that very few trades are done in the other periods during the season (only in special circumstances)).[229]

As mentioned earlier, the null hypothesis is based on Rottenberg's invariance proposition, anticipating that competitive balance should be unaffected by changes from reserve clause option systems to free agency. The original theory is based on profit maximisation, but is, according to Késenne (1996), also valid for leagues consisting of win maximising teams. The discussion below will, therefore, be related to the

[228] An expression used by Schmidt and Berri (2003).
[229] See also www.thefa.com/Features/EnglishDomestic/Postings/2002/08/22801.htm. For the general FIFA-transfer rules, see:
www.fifa.com/mm/document/affederation/administration/regulations_on_the_status_and_transfer_of_players_en_33410.pdf.

alternative hypotheses. The alternative hypotheses for the different changes in transfer market regulations in the English football league are:

1893/94 – H_1: improved competitive balance from the introduction of the retain-and-transfer system – one-sided

> This alternative hypothesis is based on two relations. The first is the general opinion (not theoretical), claiming that the reserve clause improves competitive balance. Second, as shown earlier in this chapter, Sloane (1971) argues that a number of the assumptions in Rottenberg might not be valid in European football.

1963/64 – H_1: weakened competitive balance from the change to the option-and-transfer system – one-sided

> Based on the arguments used on the alternative hypothesis above, softening the transfer market restrictions should have a negative impact on competitive balance.

1978/79 – H_1: weakened competitive balance after introducing freedom of contract – one-sided

> As for the previous alternative hypotheses, a system that increases the players' rights at the end of a contract should further weaken competitive balance.

1996/97 – H_1: improved competitive balance from the international part of Bosman I – one-sided

> As described earlier in this chapter, the first phase of the Bosman verdict had two consequences on the labour market in English football, free agency for international transfers within EU/EU-associated members and removed limitations of players from the same area. To my knowledge, the consequences of international free agency have not been analysed yet, and it is difficult to come up with an alternative hypothesis based on any direction (therefore two-sided). On the other side, the open border should increase the pool of available

talent, which is, according to the Gould hypothesis presented in Schmidt and Berri (2003), hypothesised to improve competitive balance (i.e. one-sided). This view should be even stronger, related to theoretical analyses of opening the labour market, as shown in Dobson and Goddard (2001). The combined effects should therefore lead to a one-sided alternative hypothesis – improved competitive balance.

1998/99 – H_1: weakened competitive balance as a consequence of the domestic part of the Bosman I – one-sided

The international transfer system is unchanged, while domestic transfer regulations change from player reservations to free agency, and should, on the basis of the argumentation above, lead to a one-sided alternative hypothesis - weakened competitive balance.

2002/03 – H_1: change in competitive balance from Bosman II – two-sided

The latest major change in the transfer market rule led to only minor changes in the relationship between team and player, where the most radical change (in practice) is the introduction of the transfer windows. Their effect on competitive balance are not analysed yet, and it is difficult to argue if this change will affect competitive balance in any direction.

Maximum Wage

Professional team sports have also put restrictions on teams' costs, mostly related to wages. English football first introduced absolute maximum weekly wages for (single) players in 1901. Other European football leagues had even stronger restrictions on wages, because of their amateur status, such as in Norway. In the current situation of European football, restrictions on wages and costs are much more rare, especially in the top divisions. However, this topic is popular among the popular press when teams are under financial pressure, due to relatively high wage costs. However, incentives to be competitive with other leagues, and particularly for the better teams to be able to compete in high level on UEFA team tournaments, have probably limited the use of wage restrictions in European football.

On the other side of the Atlantic, restrictions on costs have become more usual in the major leagues from the implementation of the salary cap system in the NBA at the beginning of the 1984/85 season (Fort and Quirk, 1995).[230] This salary cap has both a ceiling (maximum) and a floor (minimum), so the total wage expenditure for each team had to be in that range (see, for example, Dobson and Goddard, 2001; Fort and Quirk, 1995). According to Andreff and Staudohar (2002), the NFL followed in 1994. In the 1999-2000 season, the NBA introduced cap on individual player salaries as well. In the NHL, salary cap was introduced for the rookies in 1995 (Andreff and Staudohar, 2002; see also Staudohar, 1999).

A justification for introducing maximum wage/salary caps is that the best talent will be affordable to all teams (see, for example, Sloane, 1969; Downward and Dawson, 2000), and therefore, it will enable small-market teams to be relatively more competitive,[231] and, hence, be able to improve the distribution of playing talent in the league (Sloane, 1969; Dobson and Goddard, 2001).

The literature is confusing regarding the introduction of maximum wage in English football. Szymanski and Kuypers (1999)[232] use 1900 as the starting point for the maximum wage. Buraimo et al. (2006, p. 30) apply 1904 as the year when the FA had "its sanctioning of a maximum wage….".[233] Others use 1901 as the starting point, such as Dobson and Goddard (2001) and Morrow (1999). This thesis will follow the latter, which is consistent with Green (1953, p. 408), describing the process around the introduction of the maximum wage. It was included in the Rules of The Football Association (Rule 32) from May 1901. In other words, the maximum wage was effective at the start of the 1901/02 season in English football. It lasted until 1961, when it was removed after pressure from "the Professional Footballers Association (PFA, the

---

[230] According to Fort and Quirk, the NBA agreement was signed in April 1983.
[231] Sloane (1969) refers to the Scottish League, where there was no maximum wage. The two Glasgow teams, Rangers and Celtic, had up to that time "together won the championship on no less than 41 occasions. Yet in the case of the Football League, in 51 seasons, no first division club had won the championship more than six times, and only two clubs had achieved this record".
[232] See p. 89 and p. 251 in the 2000 edition.
[233] Szymanski and Kuypers (1999, p. 251 in the 2000 edition) claim that: "However, after some wrangling the FA finally agreed in 1904 to relinquish financial control of the game, standing back from the restrictive arrangements placed on players' transfers and wages. In many ways the FA preferred to stand above the marketplace and the League was quite happy to accept the respectability that the FA's patronage gave it."

Players' trade union)" (Dobson and Goddard (2001, p. 92).[234] From 1947, a minimum wage was also introduced, due to "the National Arbitration Tribunal". According to Green (1953), the minimum wage was for players of 20 years or older, and was £7 during the season and £5 in the "close season". At the same time, the maximum wages rose to £12 and £10 for the same periods. Sloane (1969) claims that the governing bodies in English football intended that the maximum wage system would maintain competitive balance at a higher level than otherwise. According to Sloane (1969), the consequences of removing maximum wages were that players' average wage increased substantially in the following period, but also that the number of players in the whole league were reduced by 20 % in the following six seasons. Related to competitive balance, Sloane claims that after removing the maximum wage limitation "there has been evidence of increased concentration, with clubs in large population centres becoming stronger in terms of both finances and results" (p. 186). Data from Szymanski and Kuypers (1999) and Dobson and Goddard (2001) show that: "Between 1961 and 1974 growth in wages appears to have outstripped growth in revenues among both the larger and the smaller clubs, but growth in both revenues and wages was very much faster for the larger clubs than for the smaller" (Dobson and Goddard, 2001, p 92).

The alternative hypothesis that a maximum wage system improves competitive balance, is found in Sloane (1969) as well as Fort and Quirk (1995) although the latter also includes a system of reverse-order-drafting as part of the argument. The alternative hypothesis for maximum wages is one-sided, positive for introducing and negative for removing maximum wages.

6.4.1.2. Changes in Product Market Regulations

As for many other leagues, regulations in the product market are mainly related to two relationships in English football: gate revenue sharing and the distribution of revenues from collective sale of broadcasting rights (and other rights).

[234] This is also described in Nawrat and Hutchings (1998), stating that the maximum wage system was abolished on January 9th (1961).

Revenue Sharing on Gate Revenues

At the very start of the Football League, point 8 in the body of rules stated that (Green, 1953, p. 127): "Each club shall take its own gate receipts, but shall pay its opponents a sum of £12". A fixed amount, equal to all teams, cannot be regarded as revenue sharing, since the total effect on the teams' revenues should be zero. In other words, the English Football League started without a significant gate revenue sharing system.

Gate revenue sharing became part of the sport later, and, according to Szymanski and Kuypers (1999), English football had gate revenue sharing between 1920 and 1983, based on 20 percent of a net between "minimum entrance fee fixed for each" and "the costs of staging the fixture" (p. 265). In addition, the English Football League had a "levy on gate receipts" (Morrow, 1999, p. 16), which went "to cover administrative and other joint expenses". This levy was four percent of the gate receipts. The percentage was reduced to three in 1986, and removed as the Premier League was introduced prior to the 1992/93 season (Morrow, 1999).[235]

The traditional opinion is that product market regulations should have a positive influence on competitive balance, because of better distribution of revenues among the teams in a league. However, theoretically, this view is only partly supported. For example, Fort and Quirk (1995) show that revenue sharing of gate receipts should, theoretically, have no effects on competitive balance, given profit maximising behaviour. In a league with win maximising teams, the theoretical effects are, according to Késenne (various), different, with an expected improvement in competitive balance. Since win maximising is at least as realistic as profit maximising during the history of English football (Sloane, 1971), the alternative hypothesis is that revenue sharing should be expected to have a positive effect on competitive balance in the top division of English football. In other words, one-sided tests on both introduction and removal of gate sharing (positive for introduction and negative for removing) will be applied in this chapter.

[235] This levy on gate receipt is not included in the empirical analyses in this chapter.

Revenue Sharing from Collective Sale of Broadcasting Rights

The current revenue sharing system in English football follows many other leagues in Europe, through a system where pooled revenues from sales of broadcasting rights are distributed among the teams in the league. From 1967 to 1986, all teams in the English Football (divisions one to four) shared these revenues equally (Szymanski and Kuypers, 1999; Morrow, 1999). This way of distributing revenues changed in 1986, when the First Division received 50 % of the total value, the second division 25 %, and the bottom two 12.5 % each. Within each division, the broadcasting revenues were still distributed equally among the teams (Szymanski and Kuypers, 1999). The restructuring of the English league football, when the FA Premier League broke away from the previous Football League system prior to the 1992/93 season, also included the start of a change in the sale of the broadcasting rights, as they, for the first time, were sold to "a smaller private company", the BSkyB. Interesting, from a competitive balance context, are the changes in the distribution mechanism. From this season, only a share (approximately 50 %)[236] of the revenues from the sale of the broadcasting rights is distributed uniformly among all teams in the division (FA Premier League). The rest of the revenues were further divided into two equal parts, one on the basis of sporting outcome in the given season (based on the final standing), and the other on the basis of the number of appearances on TV. This system has been unchanged ever since. Revenues from sale of the broadcasting rights on the foreign market are equally distributed among the teams (Andreff and Bourg, 2006).

The general opinion is that more concentrated distribution of the collectively earned revenues will weaken competitive balance. Késenne (2006a, p. 47) claims that "In a win-maximization league, revenue sharing always improves the competitive balance whatever the specifics of the sharing arrangement." On the other hand, Szymanski (2001, 2003c) argues that more performance related distribution of these revenues promotes competitive balance. The alternative hypothesis in this chapter follows Késenne, when assuming English football teams to be win maximising, and, hence, that more unequal distribution of revenues from sale of broadcasting rights weaken competitive balance from the 1992/93 season (one-sided).

---

[236] Szymanski and Kuypers (1999) claim that in the 1997/98 season, 45 percent of the total revenues from the broadcasting rights went to all Premier League teams, while five percent was distributed among teams that are newly relegated. Andreff and Bourg (2006) use 50 percent.

6.4.2. Structural Changes in the English Top Division

In the following, a number of changes in tournament and prize structure, which might be relevant determinants for competitive balance in English football, are presented.

The Number of Teams in the League and Relegation Procedures[237]

Already from the very beginning of the English Football League in 1888/89, the original rules included possible "punishment" for bad aggregate performance, because the bottom four teams at the end of the season had to apply for re-election (Green, 1953; Brown, 1995). According to Szymanski and Ross (2000), this system was adopted from country cricket and happened even before the league had a second division. Already in the original framework of the league, a second-class championship was mentioned (Green, 1953). As part of the creation of the Second Division from the 1892/93 season, a procedure to allow teams to move between the divisions was established, through what was called test-matches, also a term from cricket (Brown, 1995), between the bottom three in the first division and top three in the second division.[238] These were the equivalent of today's playoff matches. From the 1895/96 season, the bottom two teams in the First Division and top two teams in the Second division played test matches. According to Brown (1995), the test match system was replaced by the automatic promotion/relegation system (still bottom two from the First Division and the top two from the Second Division) from the 1898/99 season, because of lack of fairness or successfulness by the former.[239]

During the first period of the English football league, the First Division expanded several times. From the 1891/92 season, two more teams were added into the league (from 12 to 14 teams), which became 16 teams in the following season. The 1898/99 season had an expansion of two teams. From 1905/06, the First Division consisted of 20 teams, and this structure lasted until the 1919/20 season, when another two teams were

[237] In addition to the references mentioned in the text, also www.rsssf.com is used as source.

[238] See also www.rsssf.com/engpaul/FLA/1892-93.html.

[239] It is referred to the (second) test match between Stoke and Burnley at the end of the 1897/98 season, where the teams colluded to achieve a 0-0 match outcome, which gave both teams the right to play in the First Division the next season (en.wikipedia.org/wiki/The_Football_League, see also en.wikipedia.org/wiki/Football_League_Second_Division and www.rsssf.com/engpaul/FLA/1897-98.html).

added into the First Division.[240] Thereafter the tournament structure (22 teams and automatic relegation for the bottom two teams) was unchanged, until the 1973/74 season, when the number of teams automatically relegated increased by one, to three teams. This structure was unchanged until the 1986/87 season, when the fourth bottom team had to go into relegation-promotion playoff (with the bottom three teams automatically relegated). In the following season, the top division in English football was reduced to 21 teams. The relegation system from the previous season continued in the 1987/88 season and reduced the number of teams to 20 in the 1988/89 season, when the bottom three relegated automatically (only) became the standard again. In the 1990/91 season two teams were relegated (automatically), and in the following season (1991/92) the league returned to the 22 teams with the bottom three relegated automatically. This structure continued in the following two seasons, before the 1994/95 season, when the number of automatically relegated teams increased to four. From the 1995/96 season onwards the structure has remained unchanged with 20 teams in the FA Premier League, and bottom three relegated automatically to the second level division.

Under conditions of perfect competitive balance, the effects of expansions should be neutral. However, if it is possible to rank the teams regarding sporting quality, expansions of teams in a top division, with promotion and relegation to a second level tier, will mean that a higher number of relatively weaker performing teams will participate in the top division. The expansion team(s) is(are) expected to be weaker, and not average team(s). Therefore, the alternative hypothesis is the following:

H_1: Increased number of teams in the division has a negative impact on competitive balance in a promotion and relegation system – one-sided.

Studies such as Szymanski and Ross (2000) and Noll (2002)[241] conclude that relegation should, in general, have positive effects on competitive balance, when it comes to win dispersion. The main determinant is the incentive effects among the lower ranked teams to improve (also through increased financial risk-taking behaviour) the sporting quality of their squads. The hypothesis is therefore that improved win dispersion can be expected, both after the introduction of automatic relegation system in the English

[240] This expansion of teams happened with a controversy, where the last team to be included, Arsenal, was only number five in the Second Division in the previous season (1914/15) (Nawrat and Hutchings, 1998).
[241] See also Noll (2003).

football league, as well as when the number of teams to relegate increases. The effects on performance persistence might be more difficult to predict, given the way it is measured in this thesis. If the promoted teams are of a higher sporting quality than the relegating teams, one will expect the short time effect to be positive, when introducing a promotion/relegation system. Further, one might argue that the marginal effect of including one more team to promote (and relegate) is diminishing, until the team that is replaced from the top division performs better than the promoted team. The following alternative hypothesis is therefore given:

H_1: Positive effect on competitive balance when increasing the number of teams to relegate – one-sided.

Point Score System

Prior to the 1981/82 season, another structural change in English football took place, when the scoring system was reorganised from the (2,1,0) system into the "modern" (3,1,0) point score system, and hence, reduced the value of playing draw matches compared to winning.

The competitive balance measures applied in this chapter are not affected by changes in point score system. Therefore, eventually changes in the NSQF ratio, because of changes in point-score system, are connected to incentive effects among the teams. This is relevant in relation to the recent discussions in the literature (Cain and Haddock, 2006; Fort, 2007) and in Chapters 3 and 4 in this thesis about effects of measuring competitive balance in leagues where draw is a realistic outcome, such as in European football. Based on the results from the NHL, where changed point score structure might have incentive effects, the same can be related to European football, when changing from (2,1,0) to the (3,1,0) system.[242] This is supported by Haugen (2008), finding negative effects on competitive balance both theoretically and empirically (win dispersion). Following Haugen, the alternative hypothesis is:

[242] Analyses from the NHL are different from English football, for example because a given NHL team plays against teams that are not in own division or in own conference, and from incentive effects from overtime play (see, for example, Abvrevaya, 2004; Easton and Rockerbie, 2005). Easton and Rockerbie further suggest that the "new overtime rule" in the NHL should improve competitive balance through better distribution of playing talent.

H1: Negative effects on competitive balance because of changes in point score system – one-sided.

Post-Seasonal Play

One issue of high public interest in European football over the last years has been the introduction of the lucrative UEFA Champions League tournament. It allows the best teams to have a revenue source that is not available for the weaker teams, and hence, this tournament might weaken competitive balance in the domestic leagues, because of wider within-league revenue distribution.

By using qualifications for post-seasonal UEFA team tournaments as part of European football leagues' prize structures, qualifications to these tournaments can be treated more or less analogously to playoffs in the NAML, both when it comes to sporting and financial incentives.

Qualifications for these international European team tournaments have, in general, happened through sporting performance, mainly on basis of the final standings in the league. However, performance in the "FA Cup" and the "League Cup" has also led to qualification to the UEFA tournaments, as well as "fair play". Most of the rules, when it comes to the UEFA team tournaments, are decided outside the "domestic governing bodies". This chapter has made a simplification, by only focusing on the UEFA Champions Cup/Champions League,[243] partly because of the "complexity" of the qualifications for the UEFA Cup (as described above), and not least, the significance of the UEFA Champions Cup/Champions League. This is done, even though Cain and Haddock (2006), in their footnote 8 (p. 338), claim that the UEFA Cup has a greater influence on the NSQF ratio than the UEFA Champions League in regression analyses (which are not presented in their article).

Historically, qualification for the UEFA Champions Cup was limited only to winners of the domestic league championship (as well as the winners of the tournament in the previous season). Hence the European Cup only affected incentives by increasing the value of winning the domestic league championship. This increased value was both

[243] The UEFA Champions Cup (commonly known as the European Cup) was the forerunner of the UEFA Champions League.

sporting (because it allows playing in a tournament of higher prestige) and financial (because of expected[244] positive financial results from participating in the international tournament). Hence, the introduction of the UEFA Champions Cup may have had implications for domestic competitive balance. Sportingly, one would, intuitively, expect that the intensity of winning the Championship prize would be unaffected by the introduction of the Champions Cup, and hence, there was no impact on competitive balance, because the Championship was already the ultimate sporting prize to be won in the league. However, there might have been longer-run effects if participating (at least successfully) had a significant positive effect on these teams' financial situation. Given significant positive financial payoff from participating in this European team level tournament, one would expect increased investments in sporting squads by the better teams, in particular for the big market teams, who have the highest expected improvement in players' MRP.[245] The anticipated effects (which will increase with the financial payoff) on the different dimensions of competitive balance are:

1. Win dispersion might increase (especially in the long-run) because of incentives for stronger teams to further invest in improved playing squads. This effect will be higher the stronger drawing power/revenue potential those teams have. Short time effects might be related to the sporting level of strong drawing power at the time of introduction.
2. Performance persistence might first be decreased if strong market teams, due to (financial) incentives, have to "climb" on the sporting quality rankings. If the big market teams are already at the upper part of the league, even the short time effects might be rather small. On the other hand, increased risk-taking behaviour as regards expenditure on playing talent for increasing the probability of winning the prize and failing might create "yo-yo" teams,[246] who perform well but not well enough in some seasons, and then might have to spend the next number of seasons to "recover" (because of the possible financial requirement for a "fire sale" of players and then to build up a "new" team). This might lead to an increased level of volatility among some of the teams in the division.

---

[244] Since the UEFA Champions Cup was a knock out competition, the expected financial effects had a high level of uncertainty, also for teams from English football.
[245] MRP = Marginal Revenue Product. See Scully (1974).
[246] This terminology is usually used about teams moving up and down between two divisions in a merit hierarchy. Here, it is applied for teams within a certain division.

3. It should increase the probability of creating a dynasty for high drawing power teams, because the financial reward increases the probability to further invest in high sporting quality squads for the best teams.

The UEFA Champions Cup was reorganised into the UEFA Champions League in the 1992/93 season. This led to significant increased financial payoffs and also, step by step, to an increased number of teams qualifying from domestic leagues. At the time of writing this thesis, the top two teams in the English FA Premier League automatically qualify to the first round (i.e. group stage) of the UEFA Champions League. The next two teams enter the third qualifying round for the UEFA Champions League. However, the likelihood of these two English teams qualifying for the group stage is high, because of two elements. The first is a benefit from the seeding system, as a consequence of a general high sporting level among the teams from the FA Premier League that have participated in the UEFA team tournaments recently, and second, as a relatively high income league, the teams from the English FA Premier League are, in general, relatively strong competitively. At least two possible outcomes for competitive balance can be suggested from the expansion of the UEFA Champions League. First, a higher number of teams in the upper part of the sporting quality rank have the opportunity to receive an additional significant revenue source, and hence, the distribution of revenues should widen. The other is that more of the better teams can increase their probability of becoming a continuous participant in the UEFA Champions League. This is also confirmed in Hadley et al. (2005), who find that for a team participating in this season's playoff, the likelihood for this team to attend next season's playoff tripled, as the number of teams qualifying doubled in the MLB. On basis of this discussion, the effects on competitive balance, from the introduction of the UEFA Champions League, including its expansions, can be hypothesised to be:

1. When (if) the strong market teams reach the higher positions, increased probability for staying there can be expected, because of the more significant positive financial payoff from participating in this tournament. In the long run, these teams can include revenues from continuously participating in the UEFA Champions League with high probability in their budgets, and therefore be able to reach the sporting quality threshold season after season. This will increase the difference in sporting quality between these teams on the one hand, and the rest of the teams in the league on the other, and, as a consequence, widening the distribution of sporting outcome,

as is suggested in the discussions by Groot (2008). There should, therefore, be a higher probability of finding a "few teams' dynasty" at the top of the league.[247]

2. The direct consequence of this latter argument might be that it increases the performance persistence among the top teams (first tier of teams). However, higher financial rewards and a higher number of places for the UEFA Champions League qualifications may increase risk-taking behaviour among even more teams, in what could be called the second tier of teams in the FA Premier League. The consequences can be that the performance volatility from season to season, among these teams (second tier), gets higher.[248] This is because they now have increased possibility to qualify, but also that greater variance financially can lead to a higher degree of changes from season to season. In other words, the effects on performance persistence might be more dependent on where the teams are on the "sporting ranks". This means that increased performance persistence among the best teams can happen simultaneously with a higher level of volatility in the next (second) tier of teams.
3. An interesting expected consequence of increasing the number of teams to qualify for the UEFA Champion League might be that the chances of creating a "one team dynasty" have weakened in the English FA Premier League. This is because it is not only the winner that can win a high amount of prize money, but also the second, third and fourth ranked teams. Hence, the financial advantage of winning the championship, compared to the closest opponents, has been reduced. It should therefore, based on the sporting success, be more difficult for one team to outrun the others at the top of the English FA Premier League.

Another issue to take into account is that participating in a number of parallel competitions can increase the probability for injuries and limit the recovering time for players after matches, due to a tighter playing schedule. Hence, teams participating in UEFA competitions might have to invest in a higher number of high quality players, for keeping their level of sporting quality in the domestic league. These effects might slow

[247] Even if a number of these teams do not qualify for the UEFA Champions League, some of them will still get a "consolation prize" by qualifying for the UEFA cup. However, the revenues from this tournament are usually significantly lower.

[248] For example Leeds United qualified for the UEFA Champions League and got a significant increase in revenues, which were used to invest in better playing squad. However, by failing to qualify (together with other circumstances, such as "crisis" in the transfer market for football players) the additional costs were not met by the expected revenues. The team then had to sell players, which decreased the sporting level of the squad, and opened up for other teams to compete for the lucrative places (increased variation of sporting performance).

down some of the "positive" effects for these teams. as mentioned above. but can also increase the gap. Analyses of these issues are beyond the scope of this thesis.

In a perfectly balanced league, post-season play should not have any effects on competitive balance. At the beginning of each season, an equal distribution of sporting outcome is expected. However, in a real-world league, one would expect that incentive effects in the upper part of the league may increase the win dispersion. The more significant the financial value of the prize, the greater incentives, and hence, the weaker win dispersion can be expected, as well as increased probability of championship concentration. However, the latter might be reduced if the number of teams qualifying increases. The general alternative hypotheses for the different changes in the UEFA Champions Cup/League in the English top division are therefore:

H_{1A}: Introduction of the UEFA Champions Cup has a negative effect on competitive balance.
H_{1B}: Extension of the number of teams qualifying for the UEFA Champions League has a negative effect on competitive balance, except for championship concentration.

Since English teams started entering the UEFA Champion Cup, the following four periods are identified for the purposes of analysis in this chapter. These are (i) 1955 - 1996 when the winner of the league qualified (and accepted it[249]) for the UEFA Champions Cup, excluding the seasons where the teams from England were banned from these tournaments; (ii) 1996 - 1998, when the top two teams were involved in the next season's UEFA Champions League; (iii) 1998 - 2001 when this increased to three teams; and (iv) 2001 - present, when a fourth qualification spot was added. It is only in the first and the last of these four periods that no other changes have happened at the same time. The 1996/97 and 1998/99 changes coincide with transfer market changes (international Bosman and domestic Bosman). For the practical purpose of the analyses in this chapter, the last team expansion involved in the UEFA Champions League will be treated together with the 2002/03 modification of the transfer system. The alternative hypotheses for these effects on competitive balance, from these joint changes are:

[249] Descriptions about English teams to join this competition, can be found in Walvin (2000, p. 169-70). See also www.uefa.com.

1996/97: Mixed effects might be expected. where the open labour market can have a positive effect on competitive balance, while the increased number of teams to the UEFA Champions League has a rather negative effect. H_1: Effects on competitive balance – to-sided.

1998/99: The alternative hypothesis is negative, both on the labour market, due to domestic free agency, and for the play-off, because of the increased number of teams that have the possibility to qualify for the UEFA Champions League. The following H_1 is given: Negative effects on competitive balance from the 1998/99 changes – one-sided.

2002/03 (2001/02): While the increased number of teams that might qualify for the UEFA Champions League will have a negative alternative hypothesis, the alternative hypothesis for the changed rules in the transfer market is two-sided. Putting these effects together will give the following H_1: Negative effects on competitive balance for these changes jointly – one-sided.

Sporting Structure

One of the more radical changes in the playing of football happened in 1925, when the three-player offside rule was changed into a two-player offside rule (FIFA.com).[250] Walvin (2000) indicates two effects from the offside rule change that might be of interest in the context of competitive balance. First, the number of goals increased considerably, and second, as a consequence of this change, "the 'stopper' centre-half was devised. This in turn led to tactical plans to by-pass the huge men who came to dominate the centre of the pitch" (p. 138-9).[251]

Considering the number of goals in an average match, the increase is highly significant for all three periods analysed below. From the season before to the season after the introduction, the average number rose from 2.58 to 3.69, an increase of around 43 percent. Palacios-Huerta (2004) finds this to be a structural break in the time series.

[250] According to Nawrat and Hutchings (1998), the new offside rule was implemented in the English Football League from the 1925/26 season.
[251] Nawrat and Hutchings (1998) show that the Arsenal player Charlie Buchan argued for changing the tactical system, from the 2-3-5 to the 3-3-4 system.

Table 6.2: The average number of goals per match before and after changing the offside rule prior to the 1925/26 season

Offside 25/26	3-seasons Pre	3-seasons Post	5-seasons Pre	5-seasons Post	10-seasons Pre	10-seasons Post
- Mean	2.561	3.706	2.628	3.715	2.790	3.675
- Variance	0.006	0.011	0.012	0.009	0.044	0.031
- t-value	14.922***		16.892***		10.217***	
- p-value one-s	0.000		0.000		0.000	

*** Significant on one percent level.

To my knowledge, the relationships between the number of goals scored in football matches and competitive balance are not very much analysed in the literature of sport business and economics. However, a recent contribution by Groot (2008) looks at goal scoring in this context. Groot claims that (p. 34): "Perhaps the most important feature of football contributing to CB is the quite low average number of goals scored....". Further, Groot refers to Ryder (2004a, p. 7; 2004b, p. 16-17), and claims that (p. 35): "a higher scoring context gives a better winning 'resolution', while in a low scoring context, a superior team suffers more ties and losses due to bad luck than in a high scoring context". Groot claims that the reduction in the number of goals scored over time in the English top division has reduced the negative effects on competitive balance from an increased divergence in sporting quality among the teams.

Additionally, the tactical change might be linked to the competitive balance literature. As is shown in the NAML, a reason why the NBA has a weaker competitive balance than the other leagues, might be related to the special type of players, or as described in Schmidt and Berri (2003), from a "small pool of talent" (p. 698).[252] If a player from such a limited pool of talent is more significant in the play after the changed offside rule, it might weaken competitive balance, since the MRP for these players will increase, and hence, they will be more demanded by the strong drawing power teams.[253] On basis of this discussion, the alternative hypothesis for the effects on competitive balance, after the changes in the offside rule, is that: H_1: The changes in offside rules weaken competitive balance – one-sided.

[252] See also footnote 19 at page 698 in Schmidt and Berri (2003).
[253] Maximum wages and retain and transfer rules might weaken this effect, at least if the owners are maximising wins.

Other Factors

Késenne (various) shows effects on competitive balance from different team objectives. For example, competitive balance in a win maximising league is worse than in a profit maximising league. It might be that team objectives have changed during the analysed periods, and that changes might have happened differently across teams. These factors are difficult to include in the analyses of competitive balance in this chapter. The traditional view, as shown in Sloane (1971), is that sporting success is an important part of European football teams' objective function (at page 136 in Sloane), that also includes average attendance, health of the league and profit, subject to a given level of profit. However, during the 1990s, a number of teams from the FA Premier League were registered at the London Stock Exchange. One would expect these teams to be relatively more focused on financial performance (Gerrard, 2005). More recently, private owners have bought, and also removed some teams from the stock market. These team owners might have objectives that are less financially driven (at least at the team level, even if a total portfolio objective can be more related to joint profit maximisation), such as the "sportsman-owner effect" (Vrooman, 1997).[254] At the time of writing this thesis, it might very well be that the objective functions among some of the rich owners having taken over teams in England are more related to Sloane's (1971) suggestion, than pure profit maximisation.

Another question that is difficult to fully include in this study is whether commercialisation of the football market affects competitive balance. The English top division has at least gone through changes, such as described in Andreff and Staudohar (2002) about the SSSL- and MCMMG-models. The increased level of money into the game from the commercial side of football has changed the financial situation in the English top division, also for non-TV revenues. The question is whether bigger teams have a relatively higher potential on payoff from matchday and commercial revenues (i.e. sponsorship, merchandising and other venue revenues), and consequently that this will increase the revenue divergence among the teams in the league. However, by including a variable for the foundation of the FA Premier League (1992), some of the effects described above can be included (and just strengthen the one-sided alternative hypothesis about weakened competitive balance), even if Koning (2000) shows that the

[254] See Cairns et al. (1986) and Gerrard (2005) for literature review on objective functions in professional team sports.

effects of introducing shirt sponsorship in Dutch football from the 1981/82 season has not significantly affected competitive balance, but rather the general absolute quality of playing performance. On the other hand, Goossens and Késenne (2007) show theoretical effects on large market teams to be more dominant in a win maximising league after introduction of broadcasting and sponsorship revenues, when they are distributed in favour of these teams. According to Goossens and Késenne, introducing the UEFA Champions League, together with an open labour market, further increases this domination.

Other changes in sporting structure might also have effects on competitive balance. Gaviria (2000) focuses on the relationship between the popularity of more defensive play in the 1960s and negative shifts in the number of goals scored in England, Spain and Italy. The back passing rule, introduced in 1992,[255] increased the effective time the ball is in play with about ten percent (Palacios-Huerta, 2004). Analysing these changes and effects on competitive balance are beyond the limit of this thesis, as is the role of referees in the context of competitive balance (Groot, 2008).

6.4.3. Competitive Balance and Changes in Regulations and Tournament/Prize Structure in the English Top Division

In this section, competitive balance is computed for all seasons (until 2005/06) in the English top division. The NSQF ratio and the SRCC are applied to calculate competitive balance shown in figure 6.1, which also includes an overview over time for changes in regulations and tournament structure, apart from changes in the number of teams and the variations in teams to relegate in the 1980s and 1990s.

Below, figure 6.1 covers the entire sample of seasons in the top division of English football from the start in 1888/89. It contains two Y-axes. The left hand side reflects the value for the NSQF ratio, while the Y-axis on the right hand side reflects the values for the SRCC. Lower values for both measures indicate improved competitive balance.

Interpreting the overviews from the figure, it looks like the first period, after establishing the First Division, had a weaker competitive balance, while the era of

[255] See www.fifa.com/classicfootball/history/law/historylaw3.html.

maximum wages shows an improvement. Thereafter, especially after the Second World War, the general level of competitive balance seems to have become weaker, with the period after the relaxation of the labour market restrictions in the first part of the 1960s having a clear negative trend. Negative effects also seem to be the case in the period after the creation of the FA Premier League, which includes changes in the financial situation, the broadcasting exposure, the various transfer market changes, as well as the increased number of teams participating in the UEFA Champions League.

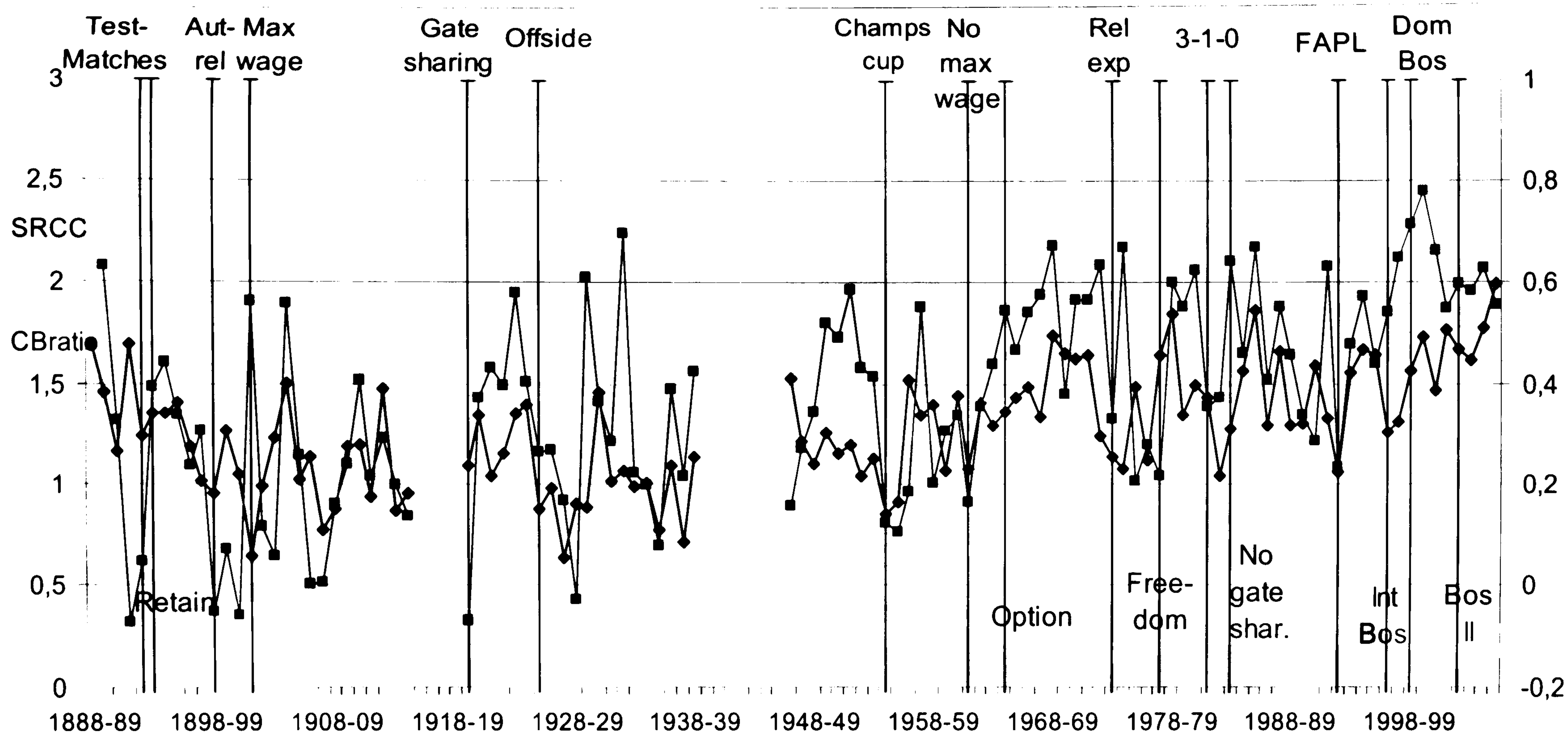

Figure 6.1: NSQF ratio and SRCC over the history of the English top division and changes in regulations and prizes[256]

[256] Both teams and schedule were repeated from the 1939/40 season, when the English Football League restarted after the SWW (www.scottishleague.net/leaguehistory.pdf). This might have influenced the trend in competitive balance just after, since the league probably would need some time to "settle", but in the system of relegation and promotion the weaker teams in the top division will be replaced by the best teams in the second level division. Therefore, one would expect that it should not take too many seasons before the "relative balance" among the teams in the division should be rebuilt. Looking at the NSQF ratio, the first season after the Second World War has the third weakest ratio (1.53). Only the first ever season of the Football League (1.69) and the fourth (1.69) had weaker ratios. The average ratio from 1888/89-19446-47 is 1.13. However, in a league with perfect competitive balance, the possible problem mentioned above should not exist, because of the anticipation that all teams in the league are of equal quality. Even newly promoted teams should be at the same level as all the other teams in the top division.

Removing revenue sharing does not seem to have negative effects on competitive balance, it looks rather positive. It might be that after a couple of seasons the performance persistence was reduced, after the increased number of teams to relegate.

There is a positive correlation between the two measures of competitive balance. The correlation coefficient for the whole sample is 0.493, while it is slightly lower in an "after Second World War" sample period (0.467). These results indicate that variations in competitive balance are tending to go in the same directions for both dispersion and persistence.

Table 6.3: Descriptive statistics on the NSQF ratio and the SRCC in the English top division

Descriptive data	NSQF ratio		SRCC	
	Whole per	After SWW	Whole per	After SWW
CB	1.282	1.407	0.375	0.460
Std dev	0.289	0.252	0.205	0.167
Minimum	0.642	0.858	-0.070	0.109
Maximum	1.992	1.992	0.777	0.777
Correlation	0.493			
Correlation		0.467		

Descriptive statistics for the two measures of competitive balance in table 6.3 show that win dispersion is wider and performance persistence is higher in the after SWW sample, indicating weakened competitive balance over time. This is consistent with Groot (2008). Moreover, for both measures, the highest level (weakest competitive balance) is found in the latest period, while the minimum points (best balanced) are found in the first part.

In this thesis, HHI has been measured, using ten season periods as base. If the same periods are applied to the NSQF ratio and the SRCC measures of competitive balance, it could be possible to compare all three dimensions of competitive balance highlighted in this thesis over time. This is done by comparing ten-seasons moving average of NSQF ratio and SRCC, with the ten-season HHI. The results, indexing the last calculation to 100 (because of easier comparison to the Norwegian league later in this chapter), from the after Second World War sample are shown graphically in figure 6.2.

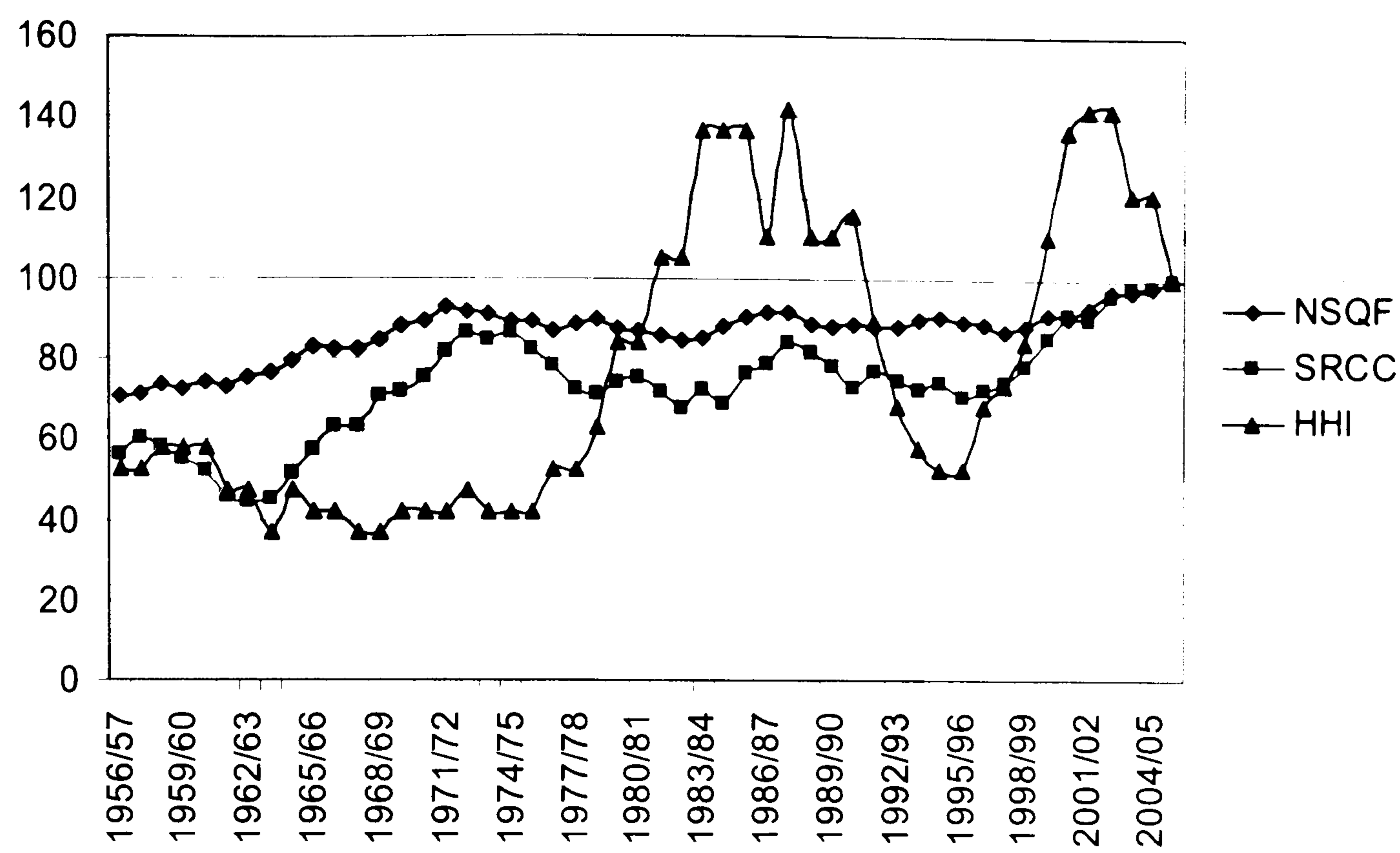

Figure 6.2: Ten-season moving average for the NSQF ratio, SRCC and HHI in the post SWW period of the English top division

For the indexed values of competitive balance, the positive coefficients for all three dimensions are all highly significant. However, there is more variation around the line for the HHI, than for the two others. Summarised, the trend after the Second World War has been towards weakened competitive balance for all three dimensions.

As is indicated in figure 6.2, the correlation between the moving average calculations for the different dimensions of competitive balance is relatively high.[257] This is particularly proved for the correlation between the NSQF ratio and SRCC, with a coefficient just above 0.9, reflecting moving average of these two dimensions to a very large degree to fluctuate in the same direction over time. The correlation between these competitive balance dimensions with championship concentration is smaller, about 0.5 (the NSQF ratio is 0.48 and the SRCC is 0.53).

The expectations that the HHI will show improvement after the extensions of the UEFA Champions League,[258] seem to be supported at the end of the sample time period for the

[257] The ten-season moving average on Kendall's τ in Groot (2008) has a relatively similar pattern as the SRCC.
[258] Compared to the Norwegian league, having a one team dynasty where the money from the UEFA Champions League participation has been an important factor for continuing the sporting performance of the particular teams, the explanation seems to be realistic.

English top division. In other words, it might be that decisions outside a domestic league's governing body, have significant relevance for both competitive balance in all the three dimensions used in this chapter, as well as for the competitive intensity at the upper part of the league.

6.4.4. Two Sample t-ratio Analysis of Competitive Balance

The next analyses follow the "first generation" tests of competitive balance effects from changes in regulations from NAML, applying simple t-ratio analyses for comparisons. Different time periods are applied. Short term effects are measured by using three seasons before and after the changes. Medium term period is reflecting the five season average competitive balance around the time for changes, while these periods are increased to ten seasons in the long term analyses. The latter analyses also include the HHI, for comparing concentration ratio for the championship winners.

6.4.4.1. Results

The results are presented in the tables below.

Table 6.4: Competitive balance and labour market changes in the English top division

	Long term						Medium term				Short term			
	NSQF ratio		SRCC		HHI		NSQF ratio		SRCC[259]		NSQF ratio		SRCC	
	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post
Retain 93/94														
- Mean							1.452	1.267	0.234	0.345	1.368	1.373	0.102	0.392
- Var							0.060	0.025	0.097	0.006	0.081	0.001	0.042	0.003
- t-value							1.414^{*}		-0.775		-0.028		-2.373^{**}	
- p-value one-s							0.097		0.232		0.490		0.038	
M Wage 01/02					0.36	0.20								
- Mean	1.256	1.061	0.166	0.238			1.100	1.083	0.102	0.311	1.094	0.961	-0.013	0.246
- Var	0.047	0.062	0.040	0.044			0.017	0.100	0.028	0.057	0.027	0.086	0.006	0.075
- t-value	1.866^{**}		-0.786				0.109		-1.609^{*}		0.684		-1.578	
- p-value one-s	0.039		0.221				0.459		0.076		0.271		0.128	
M Wage 61/62					0.22	0.16								
- Mean	1.194	1.444	0.326	0.471			1.356	1.319	0.318	0.396	1.305	1.263	0.284	0.322
- Var	0.051	0.037	0.029	0.020			0.029	0.020	0.021	0.021	0.040	0.027	0.005	0.020
- t-value	-2.661^{***}		-2.065^{**}				0.377		-0.842		0.280		-0.415	
- p-value one-s	0.008		0.027				0.358		0.212		0.397		0.353	
Option 63/64					0.18	0.18								
- Mean	1.218	1.483	0.276	0.538			1.282	1.385	0.275	0.514	1.313	1.366	0.288	0.485
- Var	0.055	0.030	0.020	0.008			0.034	0.006	0.007	0.003	0.039	0.005	0.011	0.003
- t-value	-2.884^{***}		-4.952^{***}				-1.142		-5.257^{***}		-0.439		-2.904^{**}	
- p-value one-s	0.005		0.000				0.153		0.001		0.345		0.031	
Labour start 60s					0.22	0.18								
- Mean	1.194	1.483	0.326	0.538			1.356	1.385	0.318	0.514	1.305	1.366	0.284	0.485
- Var	0.051	0.030	0.029	0.008			0.029	0.006	0.021	0.003	0.040	0.005	0.005	0.003
- t-value	-3.220^{***}		-3.514^{***}				-0.349		-2.796^{**}		-0.493		-3.895^{***}	
- p-value one-s	0.003		0.002				0.369		0.019		0.335		0.009	
Freedom 78/79					0.20	0.54								
- Mean	1.438	1.484	0.452	0.523			1.295	1.435	0.343	0.501	1.418	1.563	0.237	0.591
- Var	0.068	0.066	0.035	0.013			0.064	0.082	0.036	0.016	0.071	0.065	0.002	0.001
- t-value	-0.398		-1.019				-0.821		-1.563^{*}		-0.681		-11.244^{***}	
- p-value one-s	0.348		0.162				0.218		0.081		0.267		0.000	

[259] For the SRCC, the pre period for the calculation of the introduction of the "retain and transfer" system in 1893/94 is four seasons (the post period is five seasons).

Table 6.4 (cont.): Competitive balance and labour market changes in the English top division

	Long term[260]						Medium term[261]				Short term			
	NSQF ratio		SRCC		HHI		NSQF ratio		SRCC		NSQF ratio		SRCC	
	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post
1996/97					0.20	0.38								
- Mean	1.440	1.615	0.439	0.624			1.450	1.468	0.471	0.667	1.621	1.380	0.496	0.633
- Var	0.044	0.049	0.016	0.006			0.066	0.036	0.023	0.008	0.003	0.025	0.005	0.008
- t-value	-1.817*		-3.921***				-0.126		-2.498**		2.459*		-2.135*	
- p-value two-s	0.086		0.001				0.903		0.047		0.091		0.100	
1998/99														
- Mean	1.402	1.696	0.462	0.631			1.489	1.639	0.535	0.658	1.407	1.587	0.542	0.716
- Var	0.040	0.025	0.020	0.007			0.035	0.015	0.007	0.008	0.042	0.018	0.011	0.003
- t-value	-3.489***		-3.187***				-1.497*		-2.236**		-1.280		-2.519**	
- p-value one-s	0.002		0.003				0.089		0.028		0.145		0.043	
2002/03														
- Mean							1.569	1.760	0.669	0.589	1.656	1.683	0.661	0.601
- Var							0.035	0.028	0.007	0.001	0.027	0.006	0.013	0.001
- t-value							-1.616*		1.939*		-0.260		0.888	
- p-value one-s							0.075		0.055		0.406		0.234	

*** Significant on one percent level; ** Significant on five percent level; * Significant on ten percent level.

[260] HHI is not calculated, because the post period consists of less than ten seasons. For the NSQF ratio and the SRCC, the pre period is ten seasons, while the post period is eight seasons (1998/99).

[261] The pre period is five seasons, while the post period is four seasons (2002/03).

Table 6.5: Competitive balance and product market changes in the English top division

	Long term						Medium term				Short term			
	NSQF ratio		SRCC		HHI		NSQF ratio		SRCC		NSQF ratio		SRCC	
	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post
R sharing 19/20					0.16	0.18								
- Mean	1.047	1.083	0.194	0.280			1.091	1.202	0.252	0.344	1.104	1.163	0.211	0.247
- Var	0.044	0.059	0.015	0.042			0.063	0.021	0.011	0.059	0.109	0.026	0.006	0.075
- t-value	-0.355		-1.139				-0.853		-0.773		-0.280		-0.222	
- p-value one-s	0.363		0.135				0.209		0.231		0.397		0.418	
No RS 83/84					0.40	0.26								
- Mean	1.365	1.424	0.422	0.467			1.435	1.533	0.501	0.545	1.327	1.569	0.452	0.589
- Var	0.070	0.056	0.030	0.023			0.082	0.061	0.016	0.013	0.059	0.084	0.022	0.013
- t-value	-0.506		-0.625				-0.575		-0.581		-1.109		-1.265	
- p-value one-s	0.303		0.270				0.291		0.229		0.165		0.137	
FAPL 92/93					0.34	0.54								
- Mean	1.423	1.503	0.481	0.560			1.436	1.437	0.453	0.452	1.406	1.427	0.419	0.428
- Var	0.057	0.052	0.018	0.024			0.030	0.071	0.020	0.017	0.025	0.105	0.034	0.030
- t-value	-0.767		-1.215				-0.006		0.012		-0.101		-0.057	
- p-value one-s	0.226		0.120				0.498		0.495		0.463		0.479	

Table 6.6: Competitive balance and prize and tournament changes in the English top division

	Long term						Medium term[262]				Short term			
	NSQF ratio		SRCC		HHI		NSQF ratio		SRCC		NSQF ratio		SRCC	
	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post
Test 92/93														
- Mean							1.505	1.312			1.442	1.318	0.296	0.296
- Var							0.062	0.008			0.070	0.004	0.123	0.046
- t-value							1.629*				0.786		0.001	
- p-value one-s							0.074				0.238		0.500	
Rel 98/99					0.24	0.18								
- Mean	1.360	1.061	0.296	0.153			1.267	0.986	0.345	0.128	1.209	1.094	0.295	-0.013
- Var	0.047	0.060	0.043	0.054			0.025	0.050	0.006	0.064	0.037	0.027	0.003	0.006
- t-value	2.875***		1.411*				2.299**		1.822*		0.786		5.889***	
- p-value one-s	0.005		0.088				0.028		0.064		0.238		0.002	
Offside 25/26					0.16	0.26								
- Mean	1.167	0.986	0.297	0.307			1.263	0.861	0.438	0.258	1.308	0.838	0.460	0.236
- Var	0.047	0.042	0.033	0.044			0.024	0.017	0.007	0.053	0.017	0.031	0.010	0.003
- t-value	1.903**		-0.108				4.446***		1.646*		3.689**		3.329**	
- p-value one-s	0.037		0.458				0.001		0.069		0.011		0.015	
UEFA 55/56					0.14	0.18								
- Mean	1.166	1.285	0.378	0.321			1.080	1.250	0.410	0.272	1.014	1.259	0.324	0.283
- Var	0.029	0.038	0.023	0.024			0.018	0.061	0.030	0.029	0.020	0.096	0.030	0.056
- t-value	-1.454*		0.824				-1.347		1.263		-1.245		0.243	
- p-value one-s	0.082		0.210				0.107		0.121		0.141		0.410	
Rel 73/74					0.18	0.40								
- Mean	1.483	1.365	0.538	0.422			1.581	1.295	0.562	0.343	1.504	1.237	0.587	0.404
- Var	0.030	0.070	0.008	0.030			0.037	0.064	0.012	0.036	0.049	0.049	0.002	0.056
- t-value	1.178		1.901**				2.014**		2.242**		1.478		1.321	
- p-value one-s	0.129		0.040				0.042		0.033		0.107		0.159	
Point 81/82					0.40	0.34								
- Mean	1.405	1.434	0.468	0.454			1.491	1.439	0.455	0.500	1.563	1.257	0.591	0.458
- Var	0.067	0.056	0.034	0.016			0.076	0.093	0.036	0.021	0.065	0.038	0.001	0.025
- t-value	-0.264		0.198				0.280		-0.422		1.652*		1.420	
- p-value one-s	0.397		0.423				0.393		0.342		0.087		0.146	

*** Significant on one percent level; ** Significant on five percent level; * Significant on ten percent level.

[262] Pre period for test matches is four seasons (from the start of the league) for the NSQF ratio.

6.4.4.2. Discussions - Changes in Regulations

Labour Market Changes

Maximum Wages

A problem when analysing the effects of maximum wages on competitive balance is that of separating the effects from those of other changes occurring around the same time. The automatic relegation system was established only three seasons prior to the introduction of the maximum wage in 1901. Just two years after removing maximum wages, the transfer market restriction was modified. These events have to be taken into account when interpreting the t-ratio tests for the effects of maximum wages in competitive balance in English football. From the t-ratio tests, the short-time and medium-time effects of introducing maximum wages seem to differ between the two competitive balance methods, primarily applied in this chapter. While only small changes can be found on the dispersion of sporting outcome, performance persistence seems to have increased (as is reflected by negative t-ratio in this analysis). In other words, it does not change the relative sporting quality between the teams in general, but it increases the probability of continuing at the relative same level of ranking as before. One explanation could be that players may have less financial incentive to change teams, when there is a limit on wages, and on the other side, also when the strict transfer system is taken into account, teams may be more able to resist the financial pressures to sell players. In a free market, only the richest teams might be able to keep their best players.

The longer-term effects on win dispersion are positive and significant, while performance persistence is not significantly affected. As a consequence of the arguments above, the maximum wage (together with the retain-and-transfer system) might make it more difficult for big market teams to be dominant, and hence, narrowing the distribution of sporting output among the teams in the league. The effects on championship concentration, from introducing maximum wages, seem to be a much better distribution of championship winners.

When removing maximum wages, a similar pattern for the NSQF ratio can be found. The short term effect is rather small (with a positive sign), while the long term effect is a strongly significant weakened competitive balance. The SRCC also indicates decreased competitive balance (significant). Are these findings a result of more "power" to the bigger market teams? Note that the HHI indicates a (slightly) improved distribution of championship winners.

Transfer Systems

Retain-and-transfer system

The English Football League went from a free agency system to a player reservation system, when the retain-and-transfer system was introduced. The related t-tests are only calculated for short- and medium time periods, because these changes took place prior to the sixth season of the Football League. The calculations indicate that the change in transfer system did not have any immediate positive effect on competitive balance. Measured by the SRCC, the effect on competitive balance was actually negative over the shortest period. This result can reflect that teams could easier hold on to their better players, and hence, reduce the overall variance in rankings from season to season, at least in the short run. However, the medium term effects for the SRCC are insignificant, while the NSQF ratio shows signs of improving the competition among the teams in the league, and therefore, at a ten percent significance level, the invariance proposition is rejected.

Option-and-transfer system

As pointed out earlier, the transfer system in English football had a modification only two years after the elimination of the maximum wage system. The effects on competitive balance measured by SRCC are negative, regardless of time period, while it is only in the long-run that the negative effect can be found on the NSQF ratio.

The changes in labour market policy in the first part of the 1960s

In general, the effects of the changes in labour market policy in the early 1960s (based on analyses before the elimination of maximum wages compared to the period after the modification of the transfer system) are unconditionally negative for competitive balance measured by the SRCC, while this effect is only valid in the long run for the NSQF ratio.

Sloane (1969) claims that after the Eastham case: "The essence of the system...remains, namely that in order to obtain the services of a player with another club, a club must pay a transfer fee to the club with which the player is currently registered–though the transfer fee in the event of a dispute may be modified or quashed by either of the above-mentioned bodies" (p. 187). If the modifications of the transfer regulations, by going from the retain-and-transfer system to the option-and-transfer system, were only minor, it is the elimination of maximum wages that is the expected significant driver in relation to changes in competitive balance. Given these analyses, it seems like maximum wages (at least together with restrictive transfer market) had positive effects on competitive balance. Only one measure does not confirm this conclusion, the HHI, when removing the maximum wages.

The Michie and Oughton's index of competitive balance (C5ICB) might be used to analyse whether worsened win dispersion, as indicated by the NSQF ratio, can (partly) be explained by improved relative performance at the top of the league (top five teams). The results based on C5ICB measure in Table 6.7 follow the NSQF ratio and suggest that the long term negative influence of relaxing labour market restrictions coincides with an increased concentration at the top of the league.

Table 6.7: Labour market changes in English football at the beginning of the 1960s and the C5ICB

C5ICB	Long term		Medium term		Short term	
	1951/52-1960/61	1963/64-1972/73	1956/57-1960/61	1963/64-1967/68	1958/59-1960/61	1963/64-1965/66
Mean	1.250	1.317	1.294	1.297	1.281	1.286
Variance	0.003	0.002	0.002	0.001	0.003	0.001
t-value	-2.925***		-0.124		-0.133	
p-value one-sided	0.005		0.452		0.450	

*** Significant on one percent level.

Freedom-of-contract

The effects on competitive balance, from introducing the "freedom of contract" system, seem to be significantly negative on performance persistence in the short and medium term, but also related to the effects on the concentration of championship winners (mainly because of the Liverpool FC dominance). The effects on the NSQF ratio are insignificant for all time periods. In other words, it is only the win dispersion measure that confirms the theoretical hypothesis of changes in the transfer market (the invariance proposition). One possible explanation is that the relative differences among the teams are relatively unaffected by these changes, while the performance persistence increases significantly, especially in the shorter period of time.

Other changes in transfer market policy

The next major changes in labour market policy in English football came after the Bosman verdict. These effects are discussed together with the increased number of teams to the UEFA Champions League later in this section.

Product Market Changes

Gate sharing

Even though the negative t-value for the SRCC in the long run is higher than unit, the effects on competitive balance, from introducing revenue sharing, are rather small. These insignificant results are consistent with general theories, assuming a profit maximising regime (see, for example, Quirk and Fort, 1992; Fort and Quirk, 1995). However, they might be surprising, since in a win maximising regime, the effects are expected to be positive on competitive balance (Késenne, various).

Also when removing revenue sharing, the effects on competitive balance are insignificant, even though there might be some short time negative effects on the NSQF ratio and the SRCC, as they have t-values higher than one. The latter is expected in a league with win maximising teams. Apart from improved HHI, which is opposite of the hypothesis in a win maximising regime, no long run effect on competitive balance can be found from removing revenue sharing.

The creation of the FA Premier League (including changes in broadcasting revenue sharing systems and the general structure of the teams' revenues)

The creation of the FA Premier League, including changes in the distribution system from the collective sale of the broadcasting sales and other financial changes during this period, had no effect on competitive balance, when it comes to win dispersion and performance persistence, at least in the short-term periods. There might be (insignificant) signs of increased persistence in the long run. Moreover, the HHI indicates increased championship concentration (Manchester United domination) after the reorganisation of the league.

6.4.4.3. Discussions - Tournament and Prize Structure Changes

Test matches

The short run effects on competitive balance, from introducing test matches, are small. For the NSQF ratio, the medium term is also calculated, and supports on ten percent level the alternative hypothesis, that relegation has a positive incentive effect on competitive balance.

Relegation policy

Introducing automatic relegation in 1898/99 had positive effects on competitive balance. For the SRCC, the short-term effects are highly significant, while the long term is only weakly significant. For the NSQF ratio, positive effects are significant in medium and long term periods. These results follow, to a large degree, the suggestions in the literature about the positive effects on competitive balance after introducing a relegation/promotion system. These results also suggest that incentive effects are bigger when automatic relegation compared to test matches/relegation playoff.

Increasing the number of teams relegated in 1973/74 has also improved competitive balance, both by the NSQF and the SRCC measures. This is especially reflected in the medium time period. However, these results are not found in all dimensions of competitive balance, since the concentration of the championship winners, measured by the HHI, indicates increased concentration (Liverpool domination in the post-period). However, the latter result is not expected to be related to the number of teams to relegate, since relegation should only affect incentives at the lower part of the division. By increasing the number of teams to automatically relegate by one (= 50 %), the expected effect should be an even harder fight for survival among the bottom half teams of the division, and hence, lead to decreased NSQF ratio (improved win dispersion). On the other hand, the more teams to automatically relegate, the higher probability that the bottom team(s) will be helplessly behind, and hence having lesser incentives to play well than if fewer number of teams were to be relegated (further away from the threshold).

In addition, the improvement in competitive balance, measured by the NSQF ratio, can be a result of that the third best team in the second level division has a sporting quality that is better than the third least in the top division. If that is the case, one would expect the dispersion of sporting outcome to be narrowed in the following season, ceteris paribus. The improved SRCC might indicate that even the third best team in the second level division is performing relatively better than what the third least did before, or that the increased fear of relegation increases risk-taking behaviour for avoiding relegation. The latter might increase the volatility in performance over seasons.

Offside

The immediate effects of changing the offside rule on competitive balance seem to be positive (highly significant for both measures), totally rejecting the alternative hypothesis. In other words, the distribution of outcome narrows, and the performance variance for the teams also increases, as is the opposite of the suggested tentative hypothesis. For the longer period, the NSQF ratio is most influenced (positively) from the changes in offside rules, while the positive effects on the SRCC decrease and disappear over time. Differences between short and long run effects might be related to learning effects from the new system. The championship winner concentration is higher in the period after the changes in the offside rules, mainly because of Arsenal having a golden period in the first half of the 1930s.

Point score system

Changing the point score system seems to have a short run positive effect, both on the SRCC and NSQF ratio, while the long run effects are minimal for both measures. One would expect these changes to be related to changes in incentives, at least in the short run. A different explanation can be that the short run effects are related to uncertainty about how to play efficiently, given the new system. Learning effects will reduce these uncertainties over time, and one would therefore expect these differences to weaken in the long run, as is also confirmed by the empirical results. At least on short time, the theoretical and empirical results of increased win dispersion in Haugen (2008) are strongly rejected.

UEFA Champions Cup

Establishing the UEFA Champions Cup does not have significant effects on competitive balance, except from the long run win dispersion t-ratio analysis for the NSQF ratio. In general, if there are some effects, they are negative for the win dispersion and rather positive for the competitive balance measured by the SRCC. The HHI suggests only a minor change in championship concentration after the introduction of the post-season play.

A measure testing for increased relative sporting quality among the better teams after the championship winner also qualified for the UEFA Champions Cup, may be done by applying the Michie and Oughton's C5ICB. The pre period and the post period for this simple analysis are of equal size, and are not affected by other regulatory and regime changes. Therefore, the cut off point is the season before the maximum wages are removed.

Table 6.8: Introduction of the UEFA Champions Cup and the C5ICB

C5ICB	Pre and post UEFA Champions Cup 1949/50-1954/55	1955/56-1960/61
Mean	1.220	1.277
Variance	0.001	0.003
t-value	-2.119**	
p-value one-sided	0.030	

** Significant on five percent level.

These results support the (alternative) hypothesised effects, as the concentration in the top of the English Football League increased after the introduction of the UEFA Champions Cup.

6.4.4.4. Discussions - Combination of Transfer Market Changes and Expansions of Teams to the UEFA Champions League

1996/97 – International Bosman and two teams to the UEFA Champions League

The next major changes in labour market policy came as a consequence of the Bosman verdict, when international transfers for "out of contract" players were without financial compensations, as well as opening the labour market for players from the EU and "EU associated" countries. At the same time, the number of teams qualifying for the UEFA Champions League expanded by one for the teams in the English FA Premier League. The t-ratios show different effects on competitive balance, as a consequence of these changes, depending on time for the NSQF ratio analyses. In the long run, all measures of competitive balance go in the same direction, towards weakened competitive balance. This is consistent with Groot (2008), analysing changes in win dispersion and seasonal concentration of performance for the best team(s) over approximately the same period. However, the long run effects might be affected by other changes/modifications in this period. In the short run, competitive balance measures by the NSQF ratio improve. Theoretically, this can be explained by going from a nearly closed to a much more open labour market (see, for example, Dobson and Goddard, 2001) and the Gould hypothesis. However, this seems to be a shorter effect on the distribution of sporting outcome, but the changes in rankings from season to season decrease. The other element is the increased number (from 1 to 2) of teams in the UEFA Champions League, which might also have a negative effect on the NSQF ratio.

1998/99 – Domestic Bosman and a new expansion in the number of teams to qualify for the UEFA Champions League

The expectations for the NSQF ratio might be proved by the 1998/99 season changes, where also the shorter term effects seem to be negative, especially for the SRCC. The long-run effects are highly significant negative for both measures. In other words, the effects on competitive balance, from introducing domestic free agency and increasing the number of teams to UEFA Champions League, are negative, both for win dispersion and performance persistence, and in particular, in the long run. However, theoretically, one can say that the domestic Bosman should not affect competitive balance (invariance proposition), so it might be the UEFA Champions League effect that is the driving factor for these negative effects on competitive balance, as is argued in for example the discussions in Groot (2008).

2002/03 – Bosman II and the 2001/02 expansion of teams to UEFA Champions League

The combination of modifying the transfer system (more restricted for players under 24 and the introduction of the transfer windows), together with increased number of teams to the UEFA Champions League, had a minor effect on short time competitive balance. However, it can be found in the medium time period, when the NSQF ratio suggests a negative effect. These effects are the opposite for performance persistence.

Possible explanations for the different effects can be that further expansion of teams to the UEFA Champions League increase budgets for the best teams, and ceteris paribus, widening the gap between these teams and the weakest teams. This will increase the dispersion of sporting success, and hence, weaken competitive balance measured by the NSQF ratio. On the other side, the increased number of places for the UEFA Champions League might increase the probability that a higher number of teams are taking a more increased risk-taking behaviour for reaching top four in the league, and the teams that do not reach this goal, might weaken their potential ability the next season (getting a set back). Further, reduced persistence from season to season might appear if these high performance potential teams come back to their initial level later.

Another case to analyse is the possible effects of introducing a relatively restricted transfer windows system. On the basis of the data sample in this thesis, it is difficult to pull out effects from the transfer windows system. This is therefore a task beyond the scope of this thesis.

The interpretations from the t-test analyses of the English top division are that by weakening the labour market restrictions and increasing the number of teams to the UEFA Champions League, the distribution of sporting outcome seems to be wider and the performance persistence increases. On the other side, by increasing the rate of teams to relegate, the effects on the two dimensions of competitive balance to move in opposite directions. The distribution of championship winners seems to be relatively unaffected by these changes.

6.4.5. Time-Series Analysis of Competitive Balance in English Football

Following Maxcy (2002) and Maxcy and Mondello (2006), ordinary least square (OLS) methods in time-series data are applied to analyse the effects that regulatory regimes and structural issues have on the dependent variable, competitive balance. Competitive balance is measured by the NSQF ratio and the SRCC. These measures of competitive balance are exactly the same as Maxcy (2002) uses for the MLB, and Maxcy and Mondello (2006) apply for the other NAML.

Because of the limitations of this thesis, only one model will be used, which can be called a regime model, where changes in regulations and tournament/prize structure are specified by variables. Other models could also be relevant, such as period models, where different periods are analysed by using dummy variables, based on different time periods (more related to the t-ratio analyses above), and structural break point analysis, related to recent analyses in the NAML by Lee and Fort (2005) and Fort and Lee (2007).

The analyses for English football will be made on basis of two time periods. The first contains all seasons (1888/89-2005/06), while the other only concentrates on the after Second World War seasons (1946/47-2005/06).

Since the t- and the F-tests are based on the assumptions that the "stochastic process has a stationary (i.e. time-independent) probability distribution" (Gerrard, 2000a, p. 1), testing for stationarity for the dependent variable is necessary if the time-series regression analysis is to be classified as valid. Greene (1993) defines weakly stationarity, or covariance stationarity, as a situation where a stochastic process, y_t, has the requirements that $E[y_t]$ is independent of t, $Var[y_t]$ is a constant, independent of t, and that $Cov[y_t, y_s]$ is a function of t-s and not of t or s.

Recent research in sport economics in related topics have focused on stationarity in time series of competitive balance in NAML (e.g. Maxcy and Mondello, 2006; Lee and Fort, 2005; Fort and Lee, 2007), in addition to attendance analyses, such as in Schmidt and Berri (2001, 2002). Following Fort and Lee (2006), the first step, when using sport time series approach, should therefore be to determine if the time series are stationary or not. Consistent with their figure 1 (p. 410) about the "Time Series Approach Schematic", an ordinary unit root test is applied to suggest if the competitive balance variable for the different leagues is stationary.

Maxcy and Mondello use the Phillips-Perron test (PP) for unit roots on competitive balance in the NBA, NFL and NHL, and find that competitive balance from 1951-2004 in these leagues is characterised by stationarity, both measured by the NSQF ratio and the SRCC, by strongly significantly rejecting the unit-root hypothesis. Lee and Fort (2005) find that the unit-root hypothesis can be rejected, both for the AL and the NL in competitive balance analyses of the MLB, measured by the NSQF ratio and the tail likelihood. In addition to the PP-test, they also use the Augmented Dickey-Fuller (ADF) unit root test (with constant and time trend). The same unit-root testing methods are used in Fort and Lee (2006) for testing the same competitive balance measures in the other major leagues. Only one test does not reject the null hypothesis about unit-root, the ADF for the LTL in the NFL. The ADF tests allow for more complicated lag structures, trends and seasonal effects than the basic Dickey-Fuller test (DF), while the PP procedure tends to be applied if there are possible structural breaks. This thesis only uses the basic unit-root tests based on the DF for the different periods and measures of competitive balance (see Doornik and Hendry, 2006a). This is done because there are no indications of unit roots (see table below), and that there is no evidence of structural breaks (see, for example, Hendry and Juselius, 2000; Enders, 2004).

Table 6.9: Unit-root tests for the time series of competitive balance in the English top division

Dependent variable	Time period	t-value	Critical 5 %	Critical 1 %
NSQF	Whole period	- 5.737	- 2.889	- 3.493
	After SWW	- 4.586	- 2.911	- 3.544
SRCC	Whole period	- 6.597	- 2.889	- 3.493
	After SWW	- 5.269	- 2.912	- 3.546

The unit-root tests suggest that unit-root on all dependent variables, independent of times, can be rejected on one per cent level, and hence, anticipating stationarity for all of them. These competitive balance measures will therefore be the dependent variables in the models presented below.

The right hand side variables, except lagged dependent variables in some cases, have small degree of variation, as they are binary or are related to numbers that change seldom, such as the number of teams in the top division. Therefore, instead of testing each of those variables for unit-root, it is more appropriate to test if the linear combination of them is stationary, or that the series are co-integrated (see, for example, Hendry and Juselius, 2000; Enders, 2004). This will be done for each regression below.

Whole Period Model (1888/89-2005/06)

Independent variables:

Max wage – This is a dummy variable for the period with maximum wage – 1901/02-1960/61.

Retain – This is a dummy variable for the period when the retain-and-transfer system existed in English football – 1893/94-1962/63.

Option – This is a dummy variable for the period with the "transfer and option" transfer system – 1963/64-1977/78.

Freedom – This is a dummy variable for the period with the "freedom of contract" transfer system – 1978/79-1995/96.

Post-Bosman – This is a dummy variable for the post-Bosman period – 1996/97-2005/06.

Reg1996 – This is a dummy variable for the seasons 1996/97 and 1997/98, and it reflects three issues of relevance; the introduction of the international free agency, the opened labour market within EU/EU-associated countries, and the increased number of teams to the UEFA Champions League (from one to two teams).

Reg1998 – This is a dummy variable for the seasons 1998/99, 1999/00, 2000/01 and 2001/02. It will take into account the domestic free agency system and the increased number of teams to the UEFA Champions League (from two to three teams).

Reg2002 – This is a dummy variable for the seasons from 2002/03 and onwards, and is reflecting two changes of relevance for this thesis; the introduction of the FIFA transfer rules (that includes the transfer windows), and the increased number of teams to the UEFA Champions League (from three to four teams, which happened in the previous season - the 2001/02 season).

Gateshare – This is a dummy variable for the period with gate revenue sharing (80-20) – 1919/20-1982/83.

FAPL – This is a dummy variable for the period after the formation of the FA Premier League – 1992/93-2005/06.

Points – This is a dummy variable for the period with (3,1,0) score system – 1981/82-2005/06.

Teamnumb – This is a variable reflecting the number of teams each season.

TestPO – This is a variable reflecting the number of teams from the top division that are included in "test matches" or relegation playoff matches.

AutoREL – This is a variable reflecting the number of teams to automatically relegate.

CLparticipants – This is a dummy variable equal to 1, if the English league has team(s) to qualify for the UEFA Champions Cup or UEFA Champions League.

CLexclusion – This is a dummy variable equal to 1 for the seasons when English football teams were banned from the European football cups at team level – 19885/86-1989/90.

More about the UEFA Champions League/Champions Cup

In the analyses in this chapter, winning automatically and the qualification place for the UEFA Champions League will not be divided (i.e. 2+2 will be treated as four teams). A long tradition has been that the holder automatically qualified for the UEFA Champion Cup. Different rules apply for the UEFA Champion League, such as the situation that occurred when Liverpool won the tournament in 2005, but did not qualify automatically to defend the title the following season. UEFA recognised the anomaly and gave special dispensation for five English teams to qualify for the UEFA Champions League that season.[263] These extra teams to qualify for the tournaments are not taken into account in the analyses in this chapter.

[263] A similar situation occurred in 1999/00 involving Real Madrid (www.rsssf.com/tabless/span00.html). In that case, the Spanish football authorities operated the rule that Real Madrid would replace the fourth best team in the Spanish league in the Champions League. According to the Regulations of the UEFA Champions League 2007/08, the title holder is guaranteed a place in the group play. See Article 1 for more about effects on play in Europe for other teams in the given association (www.uefa.com/newsfiles/19071.pdf).

After Second World War Model (1946/47-2005/06)

Relevant independent variables that differ from the "whole period" model:

Max wage – This is a dummy variable for the period after removing the maximum wage limitation – 1961/62-2005/06.

Gateshare – This is a variable for the period without gate sharing (80-20) – 1983/84-2005/06

Model: $y_t = \beta_1 + \beta_2 Teamnumb_t + \beta_3 Autorel_t + \beta_4 TestPO_t + \beta_5 D_{1t} + \ldots + \beta_n D_{kt} + \varepsilon_t$, where D = dummy variable 1 to k related to the different dummy variables presented above.

Table 6.10: Time-series analysis on competitive balance and determinants in the English top division

Model	1888/89-2005/06				1946/47-2005/06			
	Post-Bosman		Post Bosm. in per.		Post-Bosman		Post Bosm. in per.	
	NSQF	SRCC	NSQF	SRCC	NSQF	SRCC	NSQF	SRCC
	(I)	(II)	(III)	(IV)	(V)	(VI)	(VII)	(VIII)
Constant	0.911[1]	-0.138	1.117[2]	-0.131	0.965	-1.535	0.965	-1.480
Std. error	(0.499)	(0.379)	(0.450)	(0.382)	(1.686)	(0.934)	(1.615)	(0.938)
Max wage	-0.084	0.110	-0.080	0.111	0.026	0.001	0.026	-0.001
	(0.118)	(0.084)	(0.114)	(0.085)	(0.183)	(0.101)	(0.175)	(0.102)
Retain	-0.141	0.177	-0.175	0.178				
	(0.182)	(0.130)	(0.173)	(0.131)				
Option	0.040	0.507[3]	0.015	0.510[3]	0.216	0.245[2]	0.216	0.274[2]
	(0.225)	(0.165)	(0.219)	(0.166)	(0.175)	(0.097)	(0.168)	(0.103)
Freedom	0.295	0.708[3]	0.276	0.711[3]	0.446[1]	0.431[3]	0.446[1]	0.461[3]
	(0.263)	(0.191)	(0.255)	(0.193)	(0.235)	(0.130)	(0.225)	(0.135)
Post-Bosman	0.433	0.888[3]			0.582[1]	0.747[3]		
	(0.276)	(0.201)			(0.297)	(0.164)		
Reg1996			0.082	0.865[3]			0.256	0.749[3]
			(0.298)	(0.225)			(0.316)	(0.187)
Reg1998			0.425	0.939[3]			0.599[2]	0.855[3]
			(0.278)	(0.214)			(0.296)	(0.184)
Reg2002			0.553[2]	0.858[3]			0.727[2]	0.756[3]
			(0.278)	(0.212)			(0.296)	(0.179)
Gateshare	0.043	0.061	0.062	0.063	0.175	0.162	0.175	0.152
	(0.118)	(0.087)	(0.114)	(0.087)	(0.202)	(0.112)	(0.193)	(0.113)
FAPL	0.198	0.045	0.205	0.046	0.106	-0.023	0.106	-0.012
	(0.151)	(0.108)	(0.146)	(0.109)	(0.170)	(0.094)	(0.163)	(0.095)
Points	-0.170	-0.106	-0.167	-0.105	-0.319	-0.224[1]	-0.319	-0.220[1]
	(0.175)	(0.125)	(0.169)	(0.126)	(0.208)	(0.115)	(0.199)	(0.116)
Clexclusion	0.340[2]	0.020	0.354[2]	0.022	0.114	-0.086	0.114	-0.122
	(0.171)	(0.119)	(0.163)	(0.120)	(0.225)	(0.126)	(0.215)	(0.128)
CLparticipants	0.148	-0.049	0.153[1]	-0.049	0.014	-0.177[2]	0.014	-0.225[3]
	(0.092)	(0.066)	(0.090)	(0.066)	(0.118)	(0.068)	(0.113)	(0.074)
Teamnumb	0.035	0.015	0.028	0.015	0.023	0.099[2]	0.023	0.101[2]
	(0.035)	(0.026)	(0.034)	(0.026)	(0.079)	(0.044)	(0.076)	(0.044)
AutoREL	-0.176[2]	-0.123[2]	-0.178[2]	-0.124[2]	-0.131	-0.102[1]	-0.131	-0.108[1]
	(0.072)	(0.051)	(0.069)	(0.051)	(0.109)	(0.060)	(0.104)	(0.061)
TestPO	-0.049	0.013	-0.057	0.013	-0.030	-0.091	-0.030	-0.087
	(0.056)	(0.041)	(0.054)	(0.041)	(0.218)	(0.121)	(0.209)	(0.121)
Offside	-0.128	0.022	-0.128	0.022				
	(0.097)	(0.069)	(0.094)	(0.069)				
Lagged dep. var.	0.026	0.042		0.033				-0.120
	(0.109)	(0.097)		(0.099)				(0.136)
R2	0.487	0.487	0.522	0.490	0.348	0.525	0.427	0.554
Observations	106	105	107	105	60	59	60	58
F	5.702[3]	5.623[3]	6.140[3]	4.916[3]	2.087[2]	4.243[3]	2.396[2]	3.474[3]
Unit-root resid	-10.06[3]	-10.85[3]	-10.32	-10.86[3]	-7.314[3]	-8.448[3]	-7.946[3]	-7.926[3]
DW	1.98	2.08	2.00	2.08	1.85	2.17	2.00	2.07
AR 1-2 test	1.261	0.802	1.135	0.775	0.617	0.264	0.420	0.209
ARCH 1-1 test	2.472	0.187	2.696	0.199	0.030	0.007	0.076	0.012
Normality test	1.795	0.963	1.218	0.916	2.113	0.552	1.350	0.618
Hetero test	0.358	1.021	0.468	0.908	0.684	0.805	0.895	0.626
RESET test	0.265	0.571	0.388	0.627	0.524	1.390	0.572	0.460

[1] Significant on ten percent level; [2] Significant on five percent level; [3] Significant on one percent level.

The results from the time-series models explain about 50 percent of the variation in performance persistence, and nearly the same for win dispersion (weaker in the SWW period) (column V and VII). Maximum wage has no significant effect on win dispersion. On performance persistence, however, maximum wage has t-value above unit when the whole sample period is applied (column II and IV). However, the sign is opposite of the hypothesised.

Following the results from the transfer market restrictions, they show that more significant variables appear for the SRCC measure than for the NSQF ratio. However, the signs and the coefficients for the latter follow the hypothesised effects, where a higher degree of restrictions gives higher coefficients. There also seems to be more negative effects on freedom-of-contract, Reg2000 and Reg2002 than the Reg1996. This might reflect positive response on opening up the football players' labour market. For performance persistence, the level of significance is much higher, and it occurs as soon as the retain-and-transfer system is modified. Coefficients and significance strongly indicate a high degree of relationship between transfer market restrictions and performance persistence. Note that the possible positive effect on competitive balance, from opening the labour market within Europe, only appears on the win distribution of performance, and not on performance persistence.

Taking the whole period into account, the theoretical hypothesis about gate sharing in a profit maximising league is supported. However, there might be some insignificant indications of negative effects when it is removed for performance persistence (column VI and VIII).

When it comes to tournament structure, an increased number of teams to be automatically relegated seem to have a significant positive influence on both measures of competitive balance. The results on the whole sample are significant, while the after SWW results have a lower degree of significance. However, t-values are still well above one for the insignificant coefficients. These results are interesting as they follow the general theoretical suggestions, as is shown earlier in this thesis.

The positive effect on competitive balance from changes in the offside rule in the t-tests earlier in this chapter is not significantly supported when taking control variables into account, even though the tendency between the two measures are upheld, as the t-value on win dispersion is greater than unity (for improved competitive balance) (column I and III).

For the whole sample period, no significant influence on the number of teams in the top division in English football can be found in relation to competitive balance, while the after SWW sample indicates significant increased persistence as the number of teams increase (column VI and VIII). The relatively frequent number of times to expand the league in the early period of the Football League, might affect the different results for the SRCC over the two samples.

Incentive effects from changing the point score system are observed in the after SWW sample, as the t-values for win dispersion are greater than one (column V and VII) and performance persistence is (weakly) significantly weaker (column VI and VIII). For the whole period sample the signs are equal, but t-values are (just) below unit. In other words, if effects from changing from the (2,1,0) to the (3,1,0) system, they are positive on competitive balance. This is not consistent with the recent theoretical development and empirical findings in Haugen (2008).

The introduction of the UEFA Champions Cup/League and the exclusion of English teams from the tournament have different results over the samples and the measures. Both variables are close to or significantly negatively related to win dispersion in the whole sample, while in the smaller sample, only SRCC has coefficients that indicate a relation; the significant positive influence on performance persistence (weaker persistence) to being involved in the UEFA Champions Cup/League.

The introduction of the FA Premier League (including changed financial situation among the teams in the league) is a highly insignificant determinant on competitive balance in the smaller sample, while in the sample covering the whole period of the top division in English football, t-values are higher than one for worsened win dispersion, but insignificant.

Analysis of why the invariance proposition might not be accepted in English football can be related to its assumptions. A difference between England and the NAML is the geographical closeness between competing teams in England, compared to the monopoly markets in major leagues cities. In England, teams are competing in partly the same markets, and this might have an effect on the revenue balance between teams. For example, there might be greater effect on incentives for competition for new prizes.

The diagnostics from the regressions show that no tests are rejected on ten percent level. Further, there is no indication on unit-root in the residuals. Except from the win dispersion in the after SWW sample (p-value is 0.037 in column V and 0.014 in column VII), the regressions have the F-test rejected on one percent level. These results are also a consequence from including a lagged dependent variable in three out of the four regressions in the whole sample period. A lagged dependent variable is also included in the SRCC after SWW sample (where the post-Bosman period is split - column VIII), which has a positive influence on the model. For the other models, this variable is not included, because it is highly insignificant.

A possible time variable is omitted from the final regressions because:

- RESET test fails for the SRCC models, when whole period is applied.
- SRCC in after SWW sample; time variable is highly insignificant (p-value 0.969 [split] and 0.842).
- NSQF ratio in the whole sample and no split in the post-Bosman period; the time variable is not included because diagnostics are rejected on ten percent level.
- NSQF ratio for the whole period and split in different regimes in the Post-Bosman period; the time variable is significant, but omitted, because it is "replacing" the significant regime dummy variables.
- NSQF ratios for the after SWW samples have the time variables insignificant (the p-value is 0.132 when one post-Bosman period in column V and 0.580 when this period is split in column VII), and they are also "replacing" the significant regime dummy variables.

6.5. Changes in Regulations and Tournament/Prize Structure in the Context of Competitive Balance. The Case of the Norwegian Top Division

The Norwegian Football Association (Norwegian FA) was founded in 1902 (Goksoyr and Olstad, 2002), joined the FIFA in 1908 (fifa.com) and has been a member of the UEFA since 1954 (uefa.com). The population in Norway is small compared to

England,[264] but the area of the country is much larger,[265] implying a much smaller population density in Norway. These conditions are important determinants, explaining why the development of the Norwegian football league, to a large degree, has lagged significantly behind developments in England. Another issue of possible relevance is the economic situation historically in Norway. When Norway became an independent country in 1905, it was one of the poorest in (Western) Europe. There are also important regional differences. For example, the Northern part of Norway was only included in the league structure from 1971 onwards (www.rsssf.no/stats/NNLeagueChamps.html).[266]

One of the characteristics of the Norwegian football league is that it retained its amateur status longer than most other European football nations until the beginning of the 1980s. The "amateur" ideal in sports was so important that, until 1969, Norwegian professional football players playing in other countries were not allowed to play for the national team (Goksøyr and Olstad, 2002, p. 315-319). It was not until the 1984 season that semi-professional player contracts could be issued. After seven years of semi-professional football, the motion for "full professionals" for players was carried in 1991,[267] and from the 1993 season all teams had to be professional to be members of the top division (through a licence, that can be called "licence of professionalism"), more than a hundred years later than in England (Goksøyr and Olstad, 2002).[268]

The professionalism of the Norwegian football league was probably an important factor in bringing more success than ever into Norwegian football, both for the national team and at team level, but also for Norwegian players abroad. For example, the national team qualified for the World Cup in 1994, for the first time since the only previous appearance in 1938. This qualification was repeated in 1998, and the qualification for Euro 2000 was the first ever. Regarding players abroad, Norwegian players were rarely represented in other leagues before the 1990s. Whereas over the whole period 1932 to 1990, only 69 Norwegian players played abroad (Goksøyr and Olstad, 2002, p. 316 and

[264] By August 2007 the population in England was just above 50 million (www.statistics.gov.uk/CCI/nugget.asp?ID=6).
[265] England: 130,395 km² (en.wikipedia.org/wiki/England) and Norway: 385,155 km² (en.wikipedia.org/wiki/Norway).
[266] These considerations are consistent with Hjelseth (2006), who also uses geography and revenue potential as two factors that might have affected the framework of the Norwegian football league. See Goksøyr and Olstad (2002) for a historical overview over Norwegian football and geography.
[267] According to Hjelseth (2006), it was implemented from the 1992 season.
[268] Additional to Goksøyr and Olstad, 2002), more about professionalism and economic development of Norwegian football can be found in for example Hjelseth (2006) and Gammelsæter and Ohr (2002).

317), while in 2001 88 Norwegian players were playing for teams outside Norway.[269] The Bosman verdict of 1995 must be taken into account, enabling free movement of football players between members and associated members of the EU (and the EEA area).[270]

At team level, the same pattern of success can be found. Apart from SK Lyn's (Oslo) success in the UEFA Champions Cup in 1968 (reaching the quarter finals), there was little success for Norwegian teams in the UEFA tournaments until 1995, when Rosenborg BK qualified for the UEFA Champions League for the first time, and then did so in nine of ten seasons thereafter. In addition, other Norwegian teams have also performed well, given the general international competitiveness of Norwegian teams in the UEFA team tournaments. For example, SK Brann (Bergen, 1996/97) and Vaalerenga (Oslo, 1998/99) both reached the quarter finals in the UEFA Cup Winners' Cups, and in 1999, Norway had two teams in the UEFA Champions League (Molde FK was the second participant).

As for the English top division, comparisons of competitive balance over time can be done by applying ten-season moving average for the NSQF ratio and for the SRCC. These measures are compared with the ten-season HHI. Because of problematic layout when using the initial calculation as an index of 100, this index is rather used in the last calculation (the 2005 ten-season moving average).

---

269 My calculation based on data from RSSSF Norwegian Football Archive (www.rsssf.no/archive.html).
270 Norway is one of the EFTA members (and non EU member) included in the European Economic Area (EEA), and has been member of the EEA from its beginning in January 1, 1994 (en.wikipedia.org/wiki/E%C3%98S).

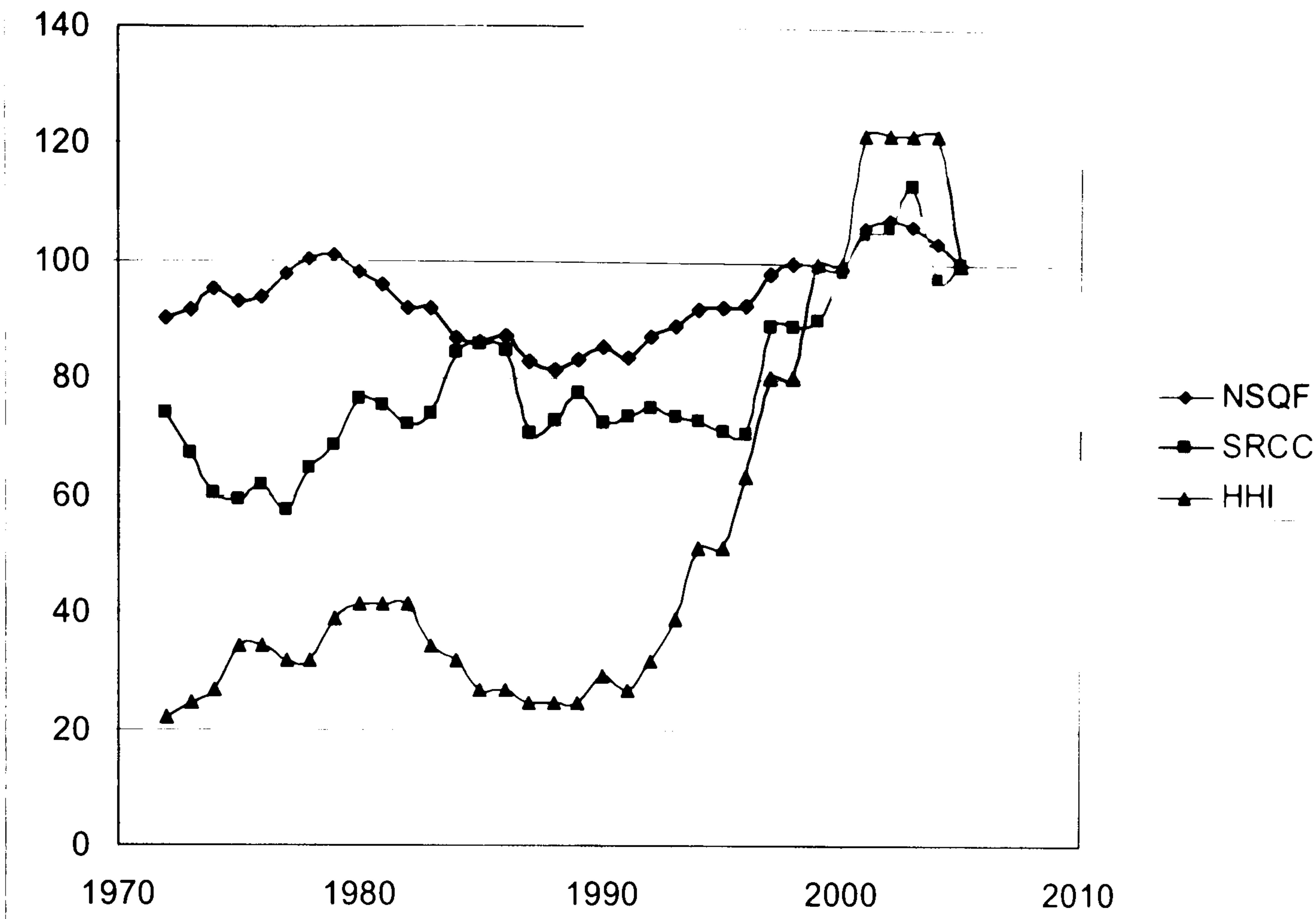

Figure 6.3: Ten-season moving average for NSQF ratio, SRCC and HHI after the restructuring of the Norwegian top division in 1963

Also for Norway, the coefficients for the three dimensions are positive, indicating weakened competitive balance over time. However, the coefficient of the NSQF ratio is lower than the English and is only significant on the five percent level. The HHI is special, because Norway had a long period of one-team dominance, where Rosenborg BK had 13 championships in a row.

Below are the correlation coefficients, based on the moving average of NSQF and SRCC, and the HHI presented:

NSQF-SRCC:	0.519
NSQF-HHI:	0.795
HHI-SRCC:	0.844

Compared to the English top division, the correlation between the ten-season moving average for the NSQF and the SRCC is much weaker in Norway. However, these measures' correlation to the HHI is higher.

6.5.1. Regulatory Changes the Norwegian Top Division

As for the English top division, the statistical null-hypothesis is the starting point (i.e. no changes in competitive balance caused by changes in regulations) in the following empirical analyses of competitive balance in the Norwegian top division.

6.5.1.1. Changes in Labour Market Regulations

Historically, transfer rules in Norwegian football were such that when a player's contract expired, there was a period of suspension if the player moved from one team to another. According to Goksøyr and Olstad (p. 301) this "suspension time" was increased from two to three months in 1950s, to four months at the end of the 1960s and further to six months in 1972. From the 1991 season, disagreements on transfer fees between the teams after a player's contract expired, were dealt with by a commission at the Norwegian FA, which settled the final transfer fee (Goksøyr and Olstad, 2002). Prior to the 1996 season, the international Bosman verdict was included in the Norwegian transfer rules. From the 1998 season, the "Bosman rule" became the rule also for transfers between Norwegian teams.[271] In the autumn of the 2002 season, the "FIFA/UEFA – EU" international transfer system, that included compensation for training young players and transfer windows, was introduced in Norwegian football (Goksøyr and Olstad, 2002).

There is little literature focusing on the relationship between competitive balance and professionalism. However, Owen and Weatherston (2002, 2004) look at this topic in New Zealand rugby. Because of differences in regulations and organisation of leagues, it is difficult to draw direct comparisons with Norwegian football. According to Owen and Weatherston, results for the period after introducing professionalism (from 1995 to 2002) suggest that competitive balance has worsened for the "National Provincial Championship" compared to the "Super 12", which is a league including teams from South Africa and Australia as well. Owen and Weatherston (2004, p. 235) claim that: "Some commentators have expressed the view that the professional game is pulling talent away from the provinces into the metropolitan unions that act as hosts to the Super 12 franchises...." Further, Owen and Weatherston (p. 238) emphasise that "a

[271] www.fotball.no/files/%7B289BD0DF-253F-4FA6-8335-74982388F68F%7D.pdf, p. 11.

degree of competitive imbalance in the NPC is not a phenomenon that is new to professionalization, and the pattern of imbalance largely reflects the distribution of the population……". More analyses are required to eventually decide if the alternative hypothesis for professionalism is one-sided. Therefore, in this chapter it will be treated as two-sided.

The alternative hypotheses in relation to competitive balance are:

H_1 (1984): Semi-professional football – two-sided.

H_1 (1991): Following the arguments from the English top division earlier in this chapter, relaxing the transfer market has negative impact on competitive balance – one-sided.

H_1 (1993): Requirement of professional football – two-sided.

H_1 (1996): Two effects because of the international Bosman verdict, where international free agency has negative effect on competitive balance and the opening of the labour market for players should have a positive effect on competitive balance. In total this suggests a two-sided alternative hypothesis.

H_1 (1998): Domestic free agency is related to negative effects on competitive balance – one-sided.

H_1 (2002): The difficulties predicting effects on competitive balance because of the changes in the regulations, lead to a two-sided alternative hypothesis.

The relationships (in time) between changes in UEFA Champions League and the changes in the labour market are much weaker for the Norwegian league, compared to the situation for the English top division. Only once during the history of the UEFA Champions League has another Norwegian team, other than the championship winner, participated in this tournament, and the dominating team during this period (Rosenborg BK) has qualified for the UEFA Champions League season after season. However, since Norway had two teams in the UEFA Champions League one season, it is not impossible that this led to risk-taking behaviour among better teams in the league. Overall, one would expect that the Post-Bosman splits will be relatively more related to the changes in the labour market restrictions for the Norwegian league, than for the FA Premier League.

6.5.1.2. Changes in Product Market Regulations

The level of gate revenue sharing has been significantly reduced over time. Until the 1989 season, the teams paid as much as 30 % of the gate revenues in sharing, where 15 % was shared with the away team, and where 15 % was a levy to the governing body. From the 1989 season, the 15 % revenue sharing to the away team was removed. In addition, the "15 % rule" to the governing body fell to five percent later on. However, the percentage to pay was then dependent on the financial solidity of the teams. Currently, it is five percent (Goksøyr and Olstad, 2002).[272]

Revenues from sale of broadcasting rights were relatively low before the 1997 deal, when the amount of NOK 135 million (about £11 million) was divided into four seasons. The next deal, covering the period 2002-2005 was NOK 300 million (TV Sports Markets, 2006). Because the amount for the teams was only a minor percent of their total revenues, specific variables related to broadcasting deals are omitted from the analyses below.

6.5.1.3. Tournament and Prize Structure Changes

The first round-robin tournament, the "Norgesserien", was also introduced a long time after the Football League in England.[273] It was only played in three seasons (1937-38, 1938-39 and 1947-48). From the 1948/49 season, the league was reorganised into the "Hovedserien", which is the starting point for analyses of competitive balance in the Norwegian top division, and is by Hjelseth (2006) claimed to be the first unitary league system in Norwegian football. This league system included two groups, where the winner of each group qualified to a double (1948/49-1950/51) or single (1951/52-1960/61) match to decide the champion. Each group had eight teams, and the bottom two teams relegated automatically. The next restructuring of the league took place prior to the 1961/62 season, when the teams in the former two groups were merged into one group of 16 teams. The championship winner was decided on basis of the winner of the final standing at the end of the season (no playoffs). The name of this league,

[272] See also Serieforeningen av 1972 (1997).
[273] Goksøyr and Olstad (2002) and rsssf.com are used as sources in this part of the chapter. The first is also a source for further overview over the creation of league systems in Norwegian football in a historical context.

"Overgangsserien ("Maratonserien")" indicates that it was only a "transitional season", before another reorganisation.

The new "First Division" was established from the 1963 season, and it consisted of ten teams, where the bottom two teams were automatically relegated. This system lasted until the 1971 season, when only one team was automatically relegated, and the top division was expanded to 12 teams from the 1972 season onwards. In addition, the changes in the tournament structure also increased the number of teams at the bottom of the league to be automatically relegated (three teams). This tournament structure was unchanged until the 1981 season, when the third last team in the league, instead of being automatically relegated, had to play a post-seasonal relegation-promotion playoff (against the second best in each of the two second level groups).

The 1986 season was the last to have the traditional (2,1,0) point score system. A (3,2,1,0) point score system was introduced for one season in the Norwegian top division. This system increased the points of winning matches to three, as well as having two outcomes for draw matches, depending on the results from the after final time shoot out, where the winner won an additional point (two in total), and the loser still got one point. As earlier, the loser of a match received zero points. In general, the "shoot-out experiment" was viewed as not very beneficial to the Norwegian top division, and was therefore removed before the start of the next season.[274] In addition, the point score system was changed again, to the current (3,1,0) system.

The next change in tournament structure in the Norwegian top division took place in the 1994 season, when only the bottom two teams were relegated automatically (with no playoff for the team finishing third bottom), as a part of the 1995 expansion to 14 teams. The relegation procedure from the 1995 season was that the bottom three teams were automatically relegated. This structure was changed from the 1997 season, when the third bottom team went into the relegation-promotion playoff system. The 1997 structure continues throughout the sample period in this thesis (including the 2005 season).

---

[274] In addition, previously the matches had been played on Sundays. In the 1987 season, one tried to copy the English tradition of playing on Saturday afternoon. This experiment was also reverted before the next season.

The alternative hypotheses for these changes are presented in the following:

H_1 (1963): Because it is difficult to predict a priori the direction of change when comparing a ten-team league with the bottom two relegated with two horizontal divisions of eight teams and two teams automatically relegated from each, a two-sided alternative hypothesis is used.

H_1 (1972): Two-sided alternative hypothesis, because expectations about improved competitive balance as the (relative) number of teams automatically relegated increases might be outweighed by an increased number of teams in the league.

H_1 (point score system changes): Following the arguments from Haugen (2008), the modern (3,1,0) system is expected to weaken competitive balance - one-sided alternative hypothesis.

H_1 (1989): Applying the same arguments as for the English league, removing revenue sharing is expected to weaken competitive balance, therefore one-sided.

H_1 (1995): Because of an increased number of teams and an increased number of automatically relegating teams (at the expense of the team to relegation-promotion playoff) can be expected to move competitive balance in opposite directions, the alternative hypothesis is two-sided in the short term period. However, since the relegation system changed back after two seasons, the alternative hypothesis in the long term period is one-sided and negative.

6.5.2. Two Sample t-ratio Analysis of Competitive Balance

Following the structure from the analyses of the English top division in football, two sample t-ratio analyses of possible changes in competitive balance, because of changes of regulations and tournament structure in Norwegian football, are presented in the following.

Table 6.11: Competitive balance and labour market changes in the Norwegian top division

	Long term						Medium term				Short term			
	NSQF ratio		SRCC		HHI		NSQF ratio		SRCC		NSQF ratio		SRCC	
	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post
Semi 84					0.28	0.32								
- Mean	1.292	1.254	0.354	0.352			1.090	1.207	0.387	0.308	1.019	1.222	0.284	0.508
- Var	0.075	0.069	0.035	0.119			0.024	0.061	0.046	0.214	0.022	0.043	0.030	0.010
- t-value	0.312		0.017				-0.892		0.349		-1.378		-1.928	
P(T<=t) two-sided	0.759		0.987				0.398		0.736		0.240		0.126	
Freedom					0.24	0.82								
- Mean	1.205	1.396	0.346	0.470			1.329	1.271	0.312	0.366	1.474	1.183	0.531	0.303
- Var	0.061	0.064	0.114	0.047			0.062	0.077	0.216	0.045	0.002	0.124	0.008	0.056
- t-value	-1.711^{*}		-0.976				0.351		-0.238		1.418		1.564^{*}	
P(T<=t) one-sided	0.052		0.172				0.367		0.409		0.115		0.096	
Prof 93					0.26	1.00								
- Mean	1.229	1.511	0.359	0.507			1.311	1.447	0.483	0.370	1.191	1.407	0.422	0.338
- Var	0.067	0.021	0.115	0.043			0.093	0.016	0.020	0.035	0.130	0.000	0.023	0.063
- t-value	-3.003^{***}		-1.179				-0.922		1.085		-1.037		0.496	
P(T<=t) two-sided	0.008		0.254				0.384		0.309		0.358		0.646	
Bos 96					0.42	0.82								
- Mean	1.300	1.406	0.339	0.478			1.271	1.521	0.366	0.574	1.407	1.573	0.338	0.451
- Var	0.063	0.083	0.117	0.069			0.077	0.027	0.045	0.034	0.000	0.039	0.063	0.005
- t-value	-0.881		-1.020				-1.731		-1.651		-1.453		-0.748	
P(T<=t) two-sided	0.390		0.322				0.122		0.137		0.220		0.496	
Bos nat 98														
- Mean	1.379	1.381	0.426	0.493			1.447	1.574	0.370	0.645	1.478	1.529	0.388	0.678
- Var	0.054	0.095	0.028	0.086			0.016	0.022	0.035	0.015	0.029	0.027	0.004	0.026
- t-value	-0.011		-0.574				-1.455^{*}		-2.756^{**}		-0.366		-2.858^{**}	
P(T<=t) one-sided	0.496		0.289				0.092		0.012		0.366		0.023	
Bos Il 02														
- Mean							1.598	1.182	0.598	0.332	1.538	1.274	0.700	0.292
- Var							0.023	0.088	0.030	0.112	0.033	0.081	0.017	0.159
- t-value							2.752^{**}		1.551		1.360		1.679	
P(T<=t) two-sided							0.028		0.165		0.245		0.168	

*** Significant on one percent level; ** Significant on five percent level; * Significant on ten percent level.

Table 6.12: Competitive balance and tournament and prize structure and product market changes in the Norwegian top division

	Long term						Medium term				Short term			
	NSQF ratio		SRCC		HHI		NSQF ratio		SRCC		NSQF ratio		SRCC	
	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post
First div 63[a]					0.36	0.18								
- Mean	1.331	1.268	0.504	0.354			1.259	1.175	0.491	0.300	1.083	1.295	0.496	0.196
- Var	0.072	0.065	0.042	0.024			0.110	0.108	0.020	0.046	0.036	0.158	0.041	0.013
- t-value	0.537		1.852*				0.405		1.658		-0.834		2.237*	
P(T<=t) two-sided	0.598		0.082				0.696		0.136		0.451		0.089	
72 restructuring					0.18	0.34								
- Mean	1.316	1.350	0.277	0.361			1.259	1.383	0.244	0.345	1.345	1.433	0.257	0.277
- Var	0.105	0.054	0.033	0.048			0.047	0.049	0.012	0.064	0.027	0.072	0.003	0.105
- t-value	-0.272		-0.933				-0.893		-0.823		-0.487		-0.110	
P(T<=t) two-sided	0.789		0.364				0.398		0.435		0.652		0.918	
Point 86-88					0.22	0.66								
- Mean	1.228	1.379	0.405	0.426			1.081	1.311	0.380	0.483	1.002	1.474	0.422	0.531
- Var	0.060	0.054	0.033	0.028			0.036	0.093	0.037	0.020	0.022	0.002	0.056	0.008
- t-value	-1.422*		-0.272				-1.435*		-0.958		-5.237***		-0.752	
P(T<=t) one-sided	0.086		0.394				0.095		0.183		0.003		0.247	
Removing RS					0.20	0.66								
- Mean	1.149	1.403	0.348	0.425			1.207	1.301	0.308	0.396	1.229	1.245	0.163	0.443
- Var	0.042	0.064	0.118	0.027			0.061	0.089	0.214	0.048	0.084	0.166	0.341	0.033
- t-value	-2.477**		-0.645				-0.544		-0.385		-0.054		-0.792	
P(T<=t) one-sided	0.012		0.263				0.301		0.355		0.480		0.236	
95 restructuring					0.42	1.00								
- Mean	1.295	1.457	0.349	0.465			1.275	1.529	0.390	0.471	1.386	1.478	0.415	0.388
- Var	0.061	0.052	0.117	0.071			0.078	0.024	0.045	0.018	0.001	0.029	0.079	0.004
- t-value	-1.529*		-0.850				-1.779		-0.715		-0.919		0.164	
P(T<=t) two-sided							0.113		0.495		0.410		0.878	
P(T<=t) one-sided	0.072		0.204				0.057		0.247					

[a] The pre-period does not include the 1961/62 season. *** Significant on one percent level; ** Significant on five percent level; * Significant on ten percent level.

Labour Market Changes

Semi-professionalism

The introduction of semi professional football in Norway had no significant effect on competitive balance, but t-values well above one, for both win dispersion and performance persistence, might indicate some negative short time tendency on competitive balance. For the longer period, the effects on competitive balance are minimal (including the HHI).

"Freedom-of-contract"

Opposite of the (one-side) alternative hypothesis, it looks like introducing the "freedom-of-contract" transfer system in 1991 had a positive short run influence on competitive balance (t-value for SRCC is significant and NSQF ratio is close to significant). However, the long-run effects seem to be more negative.

Full professionalism

Related to full professionalism in the Norwegian football league, the two most interesting results are the strong significant weakened win dispersion, when comparing ten seasons before and from the 1993 season, and the exceptional increase in the HHI. However, other factors might be more relevant determinants for these results, such as the best team's continued qualifications for the UEFA Champions League, than just the introduction of full professionalism. There is no other significant result on the other tests.

International Bosman

Introduction of the international Bosman rule might have had some negative effects on competitive balance, but these results are not significant. This period also reflects that the dominant team in the league (Rosenborg BK) had the opportunity to become even more dominant through qualification for the UEFA Champions League, and the consequent substantial increase in revenues.

Domestic Bosman

The domestic Bosman rules seem to affect the persistence most, and are significantly negative on SRCC, both on short-term and medium-term level. The latter is also significant for the NSQF ratio. Also here, the long run effects are less clear. In general, this might reflect that the domestic free agency follows the alternative hypothesis about weakened competitive balance in the Norwegian league, and, hence, that the invariance proposition is rejected.

Bosman II

The effects in the short and medium (4 seasons post period due to end of sample period) time on competitive balance in the Norwegian top division in football are positive, on both the dispersion of sporting outcome and the performance persistence dimensions, but only the NSQF in medium term is significant.

One problem, when interpreting these results, is that many of the teams in the Norwegian league went through financial problems in the first part of the current decade. This might have been partly driven by the "depression" in the international transfer market for Norwegian football players, following the "bubble" in the last part of the 1990s. Therefore, it might be that the league needed a number of seasons to "settle" again.

Product Market Changes

Gate sharing

Removing gate sharing has no effect on performance persistence in the Norwegian top division in football, but it has a long-run significant negative effect on win dispersion. The latter finding is, however, difficult to interpret, because in this period, a number of other relevant changes, in the context of competitive balance in Norwegian football, took place. Since the teams in the Norwegian league, in general, can be categorized under the objective of win maximisation, one would expect removing revenue sharing to have a negative effect on competitive balance. This is partly supported in this analysis.

Another interesting observation is the championship concentration before and after removing the gate sharing to the away team. The HHI had a big change, as it increased from 0.20 to 0.66. As explained above, it is difficult to determine if the change in the revenue sharing policy eventually has been a significant determinant for this result.

Tournament- and Prize Structure Changes

From two to one group (not including the 1961/62 season)

The changing tournament system seems to have had a positive significant effect on SRCC (weak significant in short and long term). However, because of no interdivisional play, it might be difficult to find an appropriate way of comparing competitive balance before and after the restructuring.

A reduction of the HHI indicates less championship concentration. This is inconsistent with the hypothesis, because one would anticipate that two groups with a playoff at the end of the season would reduce the probability of winning the championship for the "best team", because of increased uncertainty, when playing the one or two final deciding matches.

1972

The 1972 changes in tournament structure include an increased number of teams in the top division (from ten to twelve), as well as an increased number of teams to relegate (from two to three). The consequences of the dual tournament structure change seem to be minor. This might be a consequence of the two changes turning in the opposite directions, when it comes to predicted effects on competitive balance.

Point score policy

Note that (the significant better) competitive balance for the 1987 season, having the (3,2,1,0) point score system, is not included in this analysis. It is a comparison between the traditional (2,1,0) and the modern (3,1,0) point score systems. Changing point score system is significantly negative for the NSQF ratio. Different from the English top division, the results from the Norwegian top division support the theoretical and empirical findings in Haugen (2008) about weaker win dispersion after the point score change. Pulling out the effects from changing point score system alone might be difficult, because of the other changes that happened at about the same time, as described above. The increased HHI reflects the start of the Rosenborg BK dynasty.

1995

In short term, the 1995 restructuring from 12 to 14 teams in the Norwegian top division had minor effects on win dispersion and performance persistence. However, the long term significant negative effect on the NSQF ratio supports the alternative hypothesis about negative effects from increasing the number of teams (which might have been compensated in the short term, because of the relegation procedure changes).

<u>General comments</u>

Only a few significant results can be found in the simple t-ratio analyses. An interesting observation is that effects on competitive balance are weaker in the Norwegian top division, compared to the English.

6.5.3. Time-Series Analysis of Competitive Balance in Norwegian Football

Time-series analysis of the after Second World War time period in Norwegian football includes a number of radical changes, in both tournament structure and, not least, the professionalism of the sport. As can be shown below, unit root is rejected for both measures of competitive balance applied in the time-series analysis. The results from the unit-root tests are shown in the following table:

Table 6.13: Unit-root tests for the time series of competitive balance in the Norwegian top division

Dependent variable	Time period	t-value	Critical 5 %	Critical 1 %
NSQF	Whole period	-7.155	- 2.914	- 3.550
	After 1963	-4.868	- 2.932	- 3.593
SRCC	Whole period	-7.658	- 2.915	- 3.552
	After 1963	-6.780	- 2.932	- 3.593

The DF unit-root tests reject the null hypothesis of unit-root at one percent level for all dependent variables. As for the English top division, continuous independent variables for the Norwegian top division also have the pattern of small variation over time. Hence, unit-root tests are only done on the residuals.

Because of the diagnostics, there will be some differences in the structure of some of the independent variables for the Norwegian top division compared to the English. Instead of a continuous variable for the number of teams in the division, it is preferable to create dummy variables for Norway. Since the diagnostics in most of the Norwegian cases prefer that the number of teams to be automatically relegated and the number of teams to get a relegation playoff place are added together, this will also be done in the analysis below.

Because the expansion of teams to 14 in the Norwegian top division happened only one season before the start of the first implementation of the Bosman verdict, there will not be one dummy variable for the post-Bosman period, such as in the analysis of the English top division. In addition, the start of the (3,1,0) point score system and the elimination of the 15 percent gate sharing to the away team happened very close in time. These variables are therefore merged, to avoid multicolinearity. Further, the 1987 season is included as a control variable, because of the special point score system. For the whole sample, the same is done for the 1961/62 season, since this was a transitional season with 50 percent of the teams to be relegated.

Independent variables:

Semi-prof – This is a dummy variable for the period with semi-professionalism – 1984-1992 (using 1993 as the cut off point because the 1993 season is the first when professionalism is required in the top division).

Professional – This is a dummy variable for the period where the top division is based on full professionalism – 1993-2005.

"Freedom" – This is a dummy variable for the period after softening the transfer market until the Bosman verdict – 1991-1995.

Bosman-int – This is a dummy variable for the first period after the introduction of international (EU/EU-associated members) free agency (1996-1997), and no restriction of players from EU/EU-associated countries.

Bosman-dom – This is a dummy variable for the period after the introduction of the domestic free agency (1998-2001).

Bosman II – This is a dummy variable for the period after the Bosman II (2002-2005).

Point_gate – This is a dummy variable for the period of (3,1,0) score system and the period without gate sharing to the away team (1988-2005).

AutoPORel – This is a variable reflecting the number of teams to either automatically relegate or to relegation/promotion playoff.

1961/62 – This is a dummy variable reflecting the transitional season. It had 50 % of the 16 teams relegated, as well as it was the first season only having one single group as the top division.

1987 – This is a dummy variable for this season, having a special point score system.

Teamnumb10 – This is a dummy variable equal to unit when 10 teams in the top division and zero otherwise.

Teamnumb12 – This is a dummy variable equal to unit when 12 teams in the top division and zero otherwise.

Teamnumb14 – This is a dummy variable equal to unit when 14 teams in the top division and zero otherwise.

Model: $y_t = \beta_1 + \beta_2 AutoPORel_t + \beta_3 D_{1t} + \ldots + \beta_n D_{kt} + \varepsilon_t$, where D = dummy variable 1 to k related to the dummy variables presented above.

Table 6.14: Time-series analysis on competitive balance and determinants in the Norwegian top division

Model	1947/48-2005			
	1947/48-2005		1963-2005	
	NSQF (I)	SRCC (II)	NSQF (III)	SRCC (IV)
Constant	1.578[1]	1.020	1.365[3]	0.584[1]
Std. error	(0.833)	(0.664)	(0.408)	(0.325)
Semi-prof.	-0.071	0.161	-0.071	0.159
	(0.158)	(0.126)	(0.161)	(0.128)
Professional	0.232	0.034	0.232	0.027
	(0.309)	(0.246)	(0.315)	(0.252)
"Freedom"	-0.407[1]	-0.098	-0.407[1]	-0.105
	(0.223)	(0.179)	(0.227)	(0.182)
Bosman-int	-0.317	-0.043	-0.317	-0.040
	(0.374)	(0.298)	(0.380)	(0.304)
Bosman-nat	-0.246	0.230	-0.246	0.223
	(0.353)	(0.282)	(0.359)	(0.287)
Bosman II	-0.643[1]	-0.133	-0.643[1]	-0.129
	(0.353)	(0.281)	(0.359)	(0.287)
Point_gate	0.251	-0.010	0.251	-0.001
	(0.200)	(0.161)	(0.203)	(0.164)
AutoPORel	-0.061	-0.112	-0.061	-0.124
	(0.208)	(0.166)	(0.211)	(0.171)
1961/62	0.830	-0.197		
	(0.868)	(0.698)		
1987	-0.318	-1.006[3]	-0.318	-1.010[3]
	(0.282)	(0.225)	(0.287)	(0.230)
Teamnumb10	-0.213	-0.439		
	(0.451)	(0.360)		
Teamnumb12	-0.100	-0.267	0.113	0.187
	(0.230)	(0.184)	(0.259)	(0.212)
Teamnumb14	-0.052	-0.179	0.161	0.267
	(0.332)	(0.266)	(0.470)	(0.378)
$SRCC_{-1}$		-0.175		-0.125
		(0.120)		(0.138)
R2	0.328	0.535	0.317	0.539
Observations	57	55	43	42
F	1.616	3.282[3]	1.307	2.828[2]
Unit-root resid	-7.115[3]	-8.754[3]	-5.667[3]	-7.772[3]
DW	1.94	2.33	1.69	2.39
AR 1-2 test	0.018	1.572	1.300	1.574
ARCH 1-1 test	1.352	0.509	0.556	0.216
Normality test	0.509	0.944	0.614	1.948
Hetero test	0.415	0.621	0.464	0.477
RESET test	0.099	0.559	0.095	0.658

[1] Significant on ten percent level; [2] Significant on five percent level; [3] Significant on one percent level.

Note:

- $SRCC_{-1}$ is included to avoid rejection of diagnostics (column II and IV).
- SRCC: Trend not included because it is insignificant (both samples) or leads to inappropriate diagnostics (whole period).
- NSQF: Trend is insignificant, and it is replacing significant dummy variables (both samples).
- $NSQF_{-1}$: Insignificant variable (both samples).

Apart from the strong significant control variable for the 1987 season on performance persistence (increased variation), the models over the two different sample periods have only few variables of significance, as is also indicated by the low F-values for the NSQF. The significant F-tests for performance persistence seem to be heavily affected by the control variable mentioned above (column II and IV).

None of the other variables are significant on five percent level or better. However, on ten percent level, the Bosman II period has better win dispersion, and the freedom of contract system also seems to have a better win dispersion (weak significant on the whole sample; column I). The other variables are insignificant on win dispersion, but the changed point score system has, together with elimination of revenue sharing to the away team, a tendency to insignificantly worsen within-season competitive balance. In a reduced form on win dispersion (not presented here), these results are confirmed with the first two (Bosman II and freedom-of-contract) as significant variables, while the latter combination is significant in the whole period sample.

For performance persistence, no other variable than 1987 is significant. Because of the lack of significant variables, the results from the regime regressions are only giving some indications of variables that might affect competitive balance in the Norwegian top division, such as an insignificant negative effect on variation across seasons in the period of semi-professionalism. Also a lower number of teams have a t-value greater than unit, indicating a possible positive influence on variation of performance across seasons in the whole sample (column I). A reduced form model indicates that the latter variable might be significant, and also that implanting the Bosman verdict for domestic transfers can be a significant negative factor.

In general, there are no rejections of the diagnostics, and the unit-root tests for the residuals are rejected. However, the time-series models give minor explanations on the regime determinants for changes in competitive balance in Norwegian football. This means that these variables might not be significant explanatory factors for changes in competitive balance in the top division of football in Norway. Other alternative ways of modelling the time-series regressions can give other results, such as if applying a general-to-specific strategy for deciding the variables to be included in the final regression. This is a natural follow up for later research.

6.6. Conclusions

On basis of the analyses in this chapter, the main conclusions are listed below:

- The level of competitive balance can be affected by regulations and tournament/prize structure.
- The level of these effects and *how* they effect might vary between leagues.
- In English football, the changes in labour market restrictions from the beginning of the 1960s have had a negative impact on competitive balance (although it is difficult to separate the latest changes from the influence of the increased number of places in the UEFA Champions League).
- Changes in product market have little, if any, influence on competitive balance.
- Relegation, and particularly increases in the number of teams to be relegated, has a positive effect on competitive balance in English football. In Norway, the increased number of teams to be relegated seems to be insignificant on both NSQF and SRCC.
- Increasing the number of teams in a league has no effect on win dispersion, but there are indications of increased performance persistence, depending on the period of analysis.
- Different effects of introducing the (3,1,0) point score system are observed. In England, the effect seems to have been positive on the NSQF and significant on performance persistence (in the after SWW sample), while the effects are rather negative (non-significant) for win dispersion in Norway. However, this variable also includes elimination of revenue sharing in Norway.
- The regime determinants in the analyses have a higher explanatory degree in English football than in Norwegian football.

In the future for the two leagues, competitive balance might be most affected by changes that happen outside the leagues, since the players' labour market now is largely dependent on the UEFA, FIFA and EU, as well as the decisions made from UEFA about the number of teams qualifying for the UEFA Champions League.

The mechanisms left for the domestic governing bodies to decide, might have less influence on competitive balance, such as product market restrictions. However, effects from sharing broadcasting revenues might have more effect than these analyses have been able to pick out. Minor effects can be expected from changes in the number of teams in the league. The systems for relegation might also be sensitive on competitive balance, and can therefore be the most significant part.

7. Discussions and Conclusions

Differences in financial strength, sporting quality and sporting outcomes among contestants are typical parts of a sporting tournament's life. In this context, competitive balance and uncertainty of outcome are among the fundamentals in the business and economics of professional team sports. Tournaments with relatively weak competitive balance and uncertainty of outcome have a greater probability of offering an uninteresting product, and might therefore have higher risks related to their continued viability. This basic understanding ensures that competitive balance and uncertainty of outcome cannot be neglected by governing bodies and others interested in the well-being of a given sport or tournament.

As suggested in Cairns et al. (1986), fan support for weaker teams might fall if they are never close to competing for the league trophy. Further, Cairns et al. suggest that the interest for dominating teams can also be reduced. These relationships are further emphasised in the Report of the Commissioner's Blue Ribbon Panel on Baseball Economics (Levin et al., 2000 p. 13), which "assumes that a reasonable degree of competitive balance is an essential foundation for the continued popularity and growth of the game, and that mechanisms must be in place to ensure long-term competitive balance despite the inevitable inequalities in size, local market conditions and demographics of the communities in which MLB franchises are located". This means that also for big market teams, there is interdependency with the smaller teams to keep the league both sporting and financially viable. Without the weaker teams, the whole league might collapse. A recent example from the Italian top division in football, Serie A, emphasises these relationships. The teams in the league have individual TV-deals, and the outcome from these deals made the distribution of TV-money across the teams too unbalanced. This forced the smaller teams to take action. They refused to play and thus postponed the start of the 2002/03 season (Solberg, 2004).

Even with an academic history of more than half a century, many basic topics related to competitive balance are still ill-defined and confusing. Therefore, the first part of this thesis has its focus on conceptual issues, such as the important fundamentals of what is meant by competitive balance, how to measure it and reasons why competitive balance has been a central part of the literature of the business and economics of professional team sports. The basic view throughout the conceptual framework of this thesis is the recognition of the complexity of competitive balance. This is evident, especially when stripping the league structure to reach the core of the concept. Even in this simple league context, three fundamental dimensions of competitive balance appear: win dispersion, performance persistence and prize concentration. It is also shown that in the relatively big pool of competitive balance measures, many of them can be related to this simple league framework.

The next step in developing the understanding of competitive balance was to put the complex concept into a complex real-world league structure context. This has at least two basic consequences. First, the established simple league context measures may need to be adjusted, and, second, new measures may be required to capture the increased complexity of competitive balance. The latter is closely related to the increased number and type of prizes offered by a tournament. Post-seasonal playoff and interdivisional movement through a promotion-relegation system are two examples that are considered in this thesis.

The theoretical framework of competitive balance must be sufficiently rich to analyse a complex concept in a complex league context. This knowledge might increase the focus on the relationship between multiple prize league tournaments and competitiveness. This thesis suggests that competitiveness within a league should be related, not only to competitive balance, but also to competitions for the whole spectrum of prizes in a modern professional league. It has been argued that the focus in the relationship between competition and fan interest should be on competitive intensity related to a given tournament's prize structure.

Empirically, competitive balance has been the main issue of this thesis. The descriptive statistical focus has been related to three parts. The first is to calculate the level of competitive balance in different leagues, based on measures from the basic dimensions. Second, even if measures typically calculate how far away from, for example, perfect competitive balance a given league is, more information can be given by comparing the levels from other leagues as well. The third is related to time, because comparison is both relevant in a cross-sectional focus at a given time, and in a longitudinal setting, where trends in own league and relative comparisons with other leagues over time give meaningful information.

Competitive balance has been empirically well documented and analysed in the major leagues in North America (see, for example, Quirk and Fort, 1992). Empirical analyses are also done in European football, where the Big Five (Germany, England, Italy, Spain and France) are fully or partly involved (see, for example, Gerrard, 2004b). One aim in this thesis has been to give a broader picture of competitive balance among the member leagues of the UEFA. Empirically, this is done by using post Second World War data. A data sample of up to 60 years for a high number of leagues makes it possible to follow the suggestions mentioned above, and to compare competitive balance in many time periods, both within and between leagues in Europe, and the NAML.

The first part of comparisons among groups of leagues is done by measuring the three fundamental dimensions of competitive balance. The Big Five leagues have, on average, a better win dispersion than the other groups of leagues. At the other end of the scale, NAML have an overall weaker win dispersion than the European football leagues, while the young leagues from the former league of the Soviet Union are, on average, the weakest (last period). The latter seems to be part of a pattern, where bigger leagues, which also include Czechoslovakia and Yugoslavia, demerged into a number of smaller leagues, which again lead to weaker average win dispersion among these small leagues, compared to their "original" league. This is an interesting result that should be given more attention in future research.

Another interesting aspect in these results is the relatively wider win dispersion among the teams in the NAML, compared to European football leagues, in general. This difference has usually been explained by the fact that these are closed leagues, compared to the open European football leagues. Time-series analysis from England in this thesis confirms this hypothesis, related to the introduction of relegation. Another possible determinant for differences in the level of competitive balance across leagues might be found in the recent research by Schmidt and Berri on the Gould hypothesis, applying biology in combination with the playing structure of the sport. This aspect could have been appropriate to include in future analyses across different types of (European) team sports.

For championship concentration, at least three conclusions can be found. First, on average, there has been increased championship winner concentration over the sample period for the Big Five leagues. Second, the last decade has seen a relative high level of similarity in concentration levels across the different groups of leagues in European football. Third, concentration is much smaller in the NAML. One possible explanation for the latter result can be related to tournament structure, as the NAML all have post-season playoffs, based on knock-out systems.

Tournament structure is one of the hypothesised determinants for differences in competitive balance, and is a major focus in this thesis. However, in relation to championship winner concentration and post-seasonal championship playoff, the differences between the NAML and the European football leagues might be determined by two issues. First, European football leagues with playoffs typically arrange these over a small number of seasons, and, second, the relationship between regular season performance and the starting point in the playoffs beyond qualification is much stronger in Europe. The latter is particularly applied when all, or parts, of the regular season score are transferred into this starting point of the post-seasonal tournament (not knock-out).

The empirical performance persistence results have some of the same pattern as championship concentration. There might have been an (insignificant) increase in persistence among Big Five leagues on average, as has (significantly) also happened in small European leagues and some Eastern European leagues. On the other hand, the NAML seem to have had the opposite pattern. However, the level of persistence across groups of leagues are relatively similar in the last period analysed. Further analyses need to be done to learn more about determinants for across-seasonal changes in team performance over time, beyond what is done in this thesis. However, the time-series analysis from the English and Norwegian top divisions might give some indications. The response on performance persistence is relatively low in Norway, while for example relaxing transfer market restrictions, also in combination with an expansion of the number of teams qualifying for the UEFA Champions League, seems to have increased persistence among teams in the top division in English football. Extensions of these analyses may give better answers for why performance persistence varies less across leagues (including the NAML) than the win dispersion dimension.

Complex league system measures are applied in a number of examples, mainly related to qualifications for playoff. The results are partly driven by the measures' sensitivity to the number of teams that qualify. Treating qualification for the UEFA Champions

League similar to playoffs in NAML and Australian leagues, the concentration is, independent of measure, much higher among the Big Five leagues in Europe compared to leagues from the other continents. An interesting result from the NAML is that one of the playoff concentration measures finds that the NBA is *not* weakest. This measure finds the MLB to be the most concentrated.

Even if measuring UEFA Champions League qualification concentration is inappropriate in Norway, a comparative study is possible, using the Scandinavian Royal League as benchmark. Because of the short history of this tournament, limited data is available. However, based on four seasons, the concentration in the Norwegian football league is relatively similar to the Swedish, while the Danish league seems to have a higher concentration at the top of their league. Given that this tournament will continue, similar comparative studies should be appropriate for future research.

Another measure to be applied in a complex league structure is related to differences in sporting quality across tiers in leagues with promotion and relegation systems, which is a characteristic of European football. In particular, the pattern for the English league indicates increased differences, which might reflect the huge difference in the financial value of for example broadcasting deals between the first and second tier. However, this is not a general conclusion among the analysed leagues, because other Big Five leagues and Norway have had greater (within-league) differences in previous periods. Extensions related to the economics of relegation should be relevant topics for further research.

The closed NAML in contrast to open leagues in European football provides an important comparison of the effects of differences in prize structures. This thesis focuses on three general types of prizes in a league. These are the championship, merit qualifications and relegation. Prizes might affect competitive balance through incentive effects and are, by definition, related to competitive intensity. Tournament structure is therefore an important variable in many of the issues treated in this thesis, beyond the general empirical measures of competitive balance.

The recognition of variations in the level of competitive balance across (European) leagues motivates the empirical analyses of determinants for these differences. Unlike analyses from NAML, the relative high similarity between the European football leagues when it comes to tournament regulations, such as labour market policy, requires other factors to be taken into consideration. On the other hand, national macro factors, less relevant in similar analyses in the NAML, vary across the European football leagues. Among the hypothesised determinants applied are, for example, population, economic development and agriculture percent, together with the overall domestic sporting level. Moreover, prize and tournament structures are also included in these analyses. In other words, this thesis has not only descriptively analysed the data on competitive balance in European football, it has also focused on analysing systematic structural equalities and differences in underlying determinants for competitive balance.

The cross-sectional analyses are done by applying the three basic dimensions of competitive balance. There are differences in how these three dimensions of competitive balance react on the range of variables hypothesised to possibly affect competitive balance. The best relationship seems to be in win dispersion, where geographical, economic and sporting quality variables are significant for the average ten season dependent variable, while championship playoff replaces geography, when a five season average dependent variable is applied. Championship playoff is among the group of variables categorised into tournament structure. However, not all possible structural variables are available for analyses. The reason is that at the end of the sample period, both point score system and an elevation system through a merit hierarchy system (except from San Marino) appear in all leagues. A measure for relegation, that has been applied, is the relative number of relegated teams. Results indicate negative effects on ten season average performance persistence differences, but cannot be found as a significant determinant for differences in win dispersion. Some of the geographical variables and sporting quality might be significant drivers for differences in championship concentration. However, other factors affecting differences in championship concentration are also found to be outside the hypothesised variables analysed. One suggestion is that more "local" historical conditions within the different leagues are among these variables (see, for example, the sociology of football attendance in chapter seven in Dobson and Goddard, 2001).

The last part of this thesis attempted to complete the picture of competitive balance by using within single leagues' time-series analysis, where competitive balance is the dependent variable. These analyses show differences in the response on competitive balance from different regulatory regimes, as well as differences in tournament structure. For the two leagues, the FA Premier League in England and the top division in Norway, the hypothesised determinants seem to be more significant in English football than in the Norwegian. There are indications that relaxing labour market restrictions might weaken competitive balance, and relegation has a positive influence on competitive balance in England. An open labour market tends to have a positive effect on win dispersion. It is more difficult to find relationships in the Norwegian league, but the Bosman II regulation seems to have a positive effect of win dispersion. In the future, analyses of why the competitive balance response on the hypothesised variables exists to a much larger degree in the English top division than in the Norwegian, should be of interest, including discussions of possible determinants omitted from the analyses in this thesis.

A side analysis of competitive balance over time for the English top division might indicate that the reduced championship winner concentration at the end of the sample is something to analyse further. It appears that this reduced concentration can be related to the increased number of teams qualifying for the UEFA Champions League. One hypothesis presented in this thesis is that going from one to four qualification places not only increases the likelihood of a dominant team to continuing to qualify, but it also increases the probabilities for other teams to qualify for this prestigious and financially lucrative competition. This will, ceteris paribus, make it more difficult to hold on to sporting and financial advantages for a single dominant team. Given that players are attracted by playing in the UEFA Champions League, other teams will now increase their ability to attract such players on sporting terms, and, of course, the prospect of substantially increased revenues also makes it easier to give potential class players a competitive offer. Therefore, the increased number of teams to the UEFA Champions League might have been one explanation why Manchester United's dominant position has been reduced, as other teams from the FA Premier League have been able to qualify for this tournament without winning the domestic championship.

The higher number of qualification places for the UEFA Champions League *might* also have been attractive for rich owners to join the game (e.g. Chelsea, Aston Villa, Manchester City and West Ham). The extraordinary amount of money these owners have put into the teams has increased their teams' relative competitiveness. Chelsea FC is one obvious example. The increased number of teams qualifying for the UEFA Champions League in combination with the attractiveness of the FA Premier League might thus have increased the number of teams that are realistically able to become the championship winner. This might lead to a greater intensity at the top of the league which may challenge the dominance by the Big Four teams (i.e. Manchester United, Arsenal, Chelsea and Liverpool) in the future, and hence, weaken the performance persistence at the top of the league. This uncertainty might be extended by rules limiting constructions of a player squad in a free market, such as in the UEFA Champions League.[275] On the other hand, continued participation in the UEFA Champions League for the biggest teams in the future might increase the probability of group domination, not only because twenty percent of the teams qualify for this tournament each season, but also because new technology increasingly strengthens the scope and scale of player data.[276] As the knowledge on how to utilize this information improves, the probability of extending the accuracy in investment of expensive new players is expected to increase. If these relationships are positively related to the teams' financial capacity, this might further increase the probability of having continually high performing playing squads. The different predictions of the future at the top of the English FA Premier League could be analysed as part of a study of performance persistence among the different tiers, within the top division in English football.

The time-series analysis of the Norwegian league in which only one place has usually been available for the UEFA Champions League, is consistent with this having been one of the explanations why one team could be so dominating over many seasons. More analyses should be done in the future, as the team, Rosenborg BK, at the time of writing this thesis, seems to have lost its dominating sporting position.

A follow up demand study for Norwegian football would give an opportunity to do research on long-run uncertainty of outcome. Downward and Dawson (2000) suggest that the expected effects of the uncertainty of outcome hypothesis may be complex, by using Manchester United Football Club's success in the FA Premier League in the 1990s as an example of a dominant team where the home attendance was *not* reduced. This suggests two central questions. First, what is the criterion for being defined as a dominant team? Is the hypothesis only relevant when one team is dominant, or is it still appropriate even when two or more teams are dominating the league? The literature does not give a clear answer. Second, at what stage is the long-run dominant team

[275] See Article 17 in the Regulations of the UEFA Champions League 2007/08 (www.uefa.com/newsfiles/19071.pdf).
[276] See, for example, products by ProZone (www.pzfootball.co.uk index.htm).

defined as dominating within a particular season? Suppose that a dominant team wins the championship in the very last match of the season. This season may have had high fan interest, because the dominant team did not dominate during the whole season, and the prospect for other team(s) to end the season at the top of the table may have increased the interest in the league. Fan interest in the hope (rather than the actuality) that "underdogs" can beat the long-run dominant team/dynasties can complicate the relationship between fan interest and long-run dominance (Humphreys, 2003; Sanderson and Siegfried, 2003; Szymanski, 2003a). Humphreys (2003, p. 286) calls this "the David versus Goliath paradox". Szymanski (2003a) also points out that it might be of interest to follow the dominant team, because of the uncertainty related to whether the team will continue its dominance. Hence, domination might also have advantages for the interest/demand for the league. It can be that much focus on the dominant team at the beginning of the season (tournament), both with regards to its possibilities to continue its dominance and to other teams' possibilities to remove the dominant team from the top of the league, can increase the marketing of the league. Many of these issues question the Yankee-paradox in Vrooman (1996).[277]

Because long-run uncertainty, or in the term of Szymanski (2003b), championship uncertainty, has not been much analysed empirically, the data from the Norwegian league gives information available for future research. Additionally, including data from, for example, Scotland and England could be a good basis for analysing the possible research questions above. Applying the "glory factor", constructed by Jennett (1984), to analyse possible differences in fan behaviour between dominating team(s) and other teams when winning the championship at a relatively early stage, can be an appropriate part of such a study.

Given that what is called the Rottenberg's uncertainty of outcome hypothesis is reasonable, why has it been much more difficult to get acceptance in the empirical literature (Szymanski and Kuypers, 1999; Downward and Dawson, 2000; Dobson and Goddard, 2001; Borland and Macdonald, 2003)? One answer can be that the empirical methods, including the measures of uncertainty of outcome, have not been at an appropriate level. Another answer is the combination of multiple drivers for demand and that a general league consists of teams with unequal distribution of drawing power. These aspects question if perfect competitive balance, and hence, the highest

[277] See also Neale (1964).

uncertainty of outcome, is the optimal solution for the league aiming to maximise joint profit or joint interest, for example, measured by total attendance. A third possibility is that the hypothesis might be valid in a simple league context, but it is more difficult to find the correlation between competitive balance and attendance in a complex multi-prize league context.

The core of the importance of competitive balance is related to the interest of a sporting tournament, through the uncertainty of outcome hypothesis. This is because competitive balance is seen to be the main driver for uncertainty of outcome. In other words, the importance of competitive balance and uncertainty of outcomes is related to the fan response of a league, with regard to its competitive balance, or as emphasised in Zimbalist (2002, p. 119): "Fan perceptions and behaviour are at the core of the competitive balance problematic". However, this thesis has made a point of using an extended hypothesis related to uncertainty of outcome<u>s</u>, where the number of outcomes is related to prize structure. Therefore, both competitive balance and prize structure are hypothesised to be demand drivers. One aim in this thesis has been to improve and extend the first part of the road from competitive balance to demand. A follow-up study should, hence, apply this knowledge into demand studies.

Demand studies for football have been given for many leagues in Europe, from the first article in this area by Hart et al. (1975). The literature has, typically, focused on English football (see, for example, Bird, 1982; Walker, 1986; Peel and Thomas, 1988, 1992; Dobson and Goddard, 1992, 1995, 1996; Simmons, 1996; Baimbridge et al. 1996; Kuypers, 1997; Szymanski and Smith, 1997; Szymanski, 2001; Forrest and Simmons, 2002). There are also studies of football in Germany (Czarnitzki and Stadtmann, 2002) and the former West-Germany (following the team Hamburger SV) (Gärtner and Pommerehne, 1978), in Belgium (Janssens and Késenne, 1987), and Spain (Garcia and Rodriguez, 2002). The demand for Scottish football has been analysed in a number of studies (Jennett, 1984; Cairns, 1987; Smart and Goddard, 1991; Peel and Thomas, 1996). As a Norwegian, a natural follow up to this thesis is a demand study for Norwegian football.

Sloane (1971) claims that even if teams' objective is to obtain the highest possible sporting level, they are also interested in success of their rivals. This is because "the more successful the rival in terms of league position and popularity, the larger will be the total attendance resulting from the common product" (p. 124). This follows a general anticipation about the relationship between the demand and sporting success of a given team, the demand increases as the sporting success improves, but at a given point of sporting success, the team will be too dominating, and hence stronger domination will lower the interest for this team. However, future research should not only focus on this general relationship, but also on possible segmentations on the demand-side for sport.

Cairns et al. (1986) claim that consumers of professional team sports can be divided into two groups. The first is the spectators attending the match or following the match on radio and TV (and now also on internet). The other group is firms through, for example, sponsorship and advertising. These two main groups might be called household supporters and business-to-business consumers.[278] Concentrating on household supporters, these can further be divided into two parts, where one can be called "team supporters", and the other "match or league supporters" (see also Gerrard, 2006b). A match-/league-supporter is a supporter who cares about the sport, but does not have any particular preferences for the teams participating in a particular match. Related to domination, interpreting Cairns et al. (1986), team supporters are further split into two groups. The first is team fans, valuing the fact that their team is in a close championship race. Second, individual consumers (team fans) appreciate success for their team, but not valuing a close contest. In other words, the first group of team fans will indirectly (and ironically) receive utility from the success of other teams (in the battle for the championship). The other type of team fans' utility curve should be strictly increasing with improved sporting success of their team, even when the team is the dominating one. Further, the marketing literature has a broader separation of fans. According to Shank (2002), Sutton et al. (1997) use three categories of fan identification. First, there are low identification fans attending due to social or entertainment reasons. Medium identification fans, or focused fans, are fans that identify with a team or player for a short period of time, but mainly when success is achieved. Of course, for both these groups of fans, it is possible that fans "transfer" to a higher level. The high identification fans are fans that have "a long-term commitment to the sport, team, or player" (Shank, p. 208).

---

[278] Shank (2002) calls them individual and corporate consumers.

The relationship between the different segments of spectators and their understanding and response on competitive balance and uncertainty of outcomes are topics for a follow up to this thesis. It could be applied in a cost-benefit analysis related to fan reactions and implications on revenues and might require a combination of both quantitative and qualitative research methods.

The introduction of the concept of competitive intensity is an innovative part of this thesis and has, during the process, created a new principal arm in the "tree of knowledge" of the competitive balance area. Competitive intensity will be affected from both the intensity of the competition for prizes and the number of significant prizes in the tournament. Further analyses of the concept competitive intensity should follow natural time structures, as can be found in both uncertainty of outcome, competitive balance and demand studies. The suggestion is therefore to divide the concept into *end-of-season competitive intensity* and *within-season competitive intensity*. End-of-season competitive intensity would capture the intensity for prizes on basis of analyses from the final (end-of-season) league table. This means that end-of-season competitive intensity and competitive balance can be compared, as well as being part of demand studies related on seasonal data. However, future research depends upon creating appropriate measures for end-of-season competitive intensity.

Analyses of within-season competitive intensity will be even more complicated and complex. It should focus on the match significance literature, and, hence, should follow the thoughts by Jennett (1984). Match-level data can capture the fluctuation in intensity during the season much better than an end-of-season data and can therefore be applied in cross-sectional and panel data demand studies. Development of within-season competitive intensity should therefore be a prioritized task for future research.

In European football, the cross-border UEFA Champions League is the highest quality league. Two directions of further research seem obvious. One is to extend the analysis of the impact of the UEFA Champions League in the domestic leagues, beyond the effects on competitive balance, as is analysed in this thesis. In sporting terms, one will expect the introduction and extension of this tournament to have an impact on competitive intensity in leagues. Additionally, financial impacts from the expected financial payoffs, from qualifying for this league, are also highly significant. A second topic is to analyse competitive balance in the Champions League itself, as a follow up to Gerrard (2004b).

Many issues related to this thesis can be used in analysing consequences of a possible future introduction of a European Super League in football. Even if Hoehn and Szymanski (1999) have done such analyses, other topics, such as for example effects on competitive intensity in the domestic leagues, from eventually changing the UEFA Champions League to a closed European Super League, should be of relevance in future research.

Other, perhaps more realistic, changes in tournament structure in European football can also be analysed, such as changes in the number of teams to qualify for the UEFA Champions League from the top leagues.[279] Will this imply improved win dispersion, weakened performance persistence, but higher championship concentration in the FA Premier League? This could be the analogous interpretation from the results on competitive balance related to increasing the number of UEFA Champion League places in a domestic league. Moreover, similar time-series analysis could also be done for other European football leagues, to find a clearer picture on consequences on competitive balance in domestic leagues.

The Scandinavian Royal League was introduced in 2004. This league is a new phenomenon, where separate domestic leagues go together and create an international playoff for their top teams. The Royal League is accepted by the UEFA, but does not have any relations with the UEFA tournaments (i.e. there is no qualification for UEFA tournaments). Extended analyses of possible consequences for domestic league competitive balance (which is also a part of this thesis) and the relationship to competitive intensity for the domestic leagues, are possible research areas for the future. This knowledge might be even more relevant, if similar leagues appear in the future.[280]

Analyses of competitiveness within multiple prize competitions are not only relevant in professional team sports. Most individualistic sports can also be put into this categorisation. The influence of "sub-" competitions (prizes) within a greater

[279] This is for example suggested by the UEFA President Michel Platini. See www.timesonline.co.uk/tol/sport/football/european_football/article2356084.ece.

[280] The Baltic League is a similar tournament. See www.balticleague.com/standings/2007/final/. Setanta Sports Cup is an All-Ireland Cross-Border cup. See www.rsssf.com/tablesa/alliercuphist.html and en.wikipedia.org/wiki/Setanta_Sports_Cup.

competition, such as Tour de France, "Vasaloppet"[281] and the new competition in cross-country skiing called Tour de Ski,[282] should allow further scope to investigate the relevance of applying competitive intensity beyond professional team sports.

The recent contribution on competitive balance in European football by Groot (2008) includes discussions about policy implications on basis of the negative trends in competitive balance among the Big Five leagues, except the Spanish league, as well as the Dutch league. In particular the post-Bosman verdict period (that also includes changes in the UEFA Champions League) seems to have been negative on win dispersion and concentration in sporting output at the top of the leagues (see table 6.1 in Groot).

Groot looks to other contributors to possibly solve within-league competitive balance problems, such as Hoehn and Szymanski (1999) about forming a European Super League, a NAML look-a-like league, and to Késenne (2005b). Following Késenne (2007c), a problem in the post-Bosman and UEFA Champions League period is the deviation between a closed product market (domestic leagues) and an open labour market with the consequence that the differences between teams from different leagues have escalated. However, the theoretical analysis does not claim that this should weaken the domestic competitive balance. Therefore Késenne's argument for an international European Super League is based on competitive balance problems between leagues. Groot suggests a different approach to solve the current "competitive balance problem", by turning the clock backwards to the systems from the 1950s with matches shown on free-to-air television without commercials (others than covering the costs). In other words, Groot wants to welcome a system close to the Spectators-Subsidies-Sponsors-Local model in Andreff and Staudohar (2002), but, contrary to the others mentioned above, within the current dual competition system including both domestic and international competitions.

This thesis has emphasised that it is important not to over focus one dimension of competitive balance, since it is a multi-dimensional concept, where all dimensions are pieces in the broad picture of competitive balance. It has also been a focus on within-league competitive balance, both with regards to level and determinants. Analysing

---

[281] www.vasaloppet.se.

[282] www.tour-de-ski.com .

between-league competitive balance is, however, outside the framework of this thesis. Looking at within-league competitive balance, there might still be a question if eventually competitive balance problems are big enough to reconstruct the whole system of European football in this context. Except from the average of the new leagues in the former Soviet Union area, win dispersion is still (last decade of the analysis) on average better in European football, and this is in particular the case for the Big Five leagues, compared to the NAML. Performance persistence is better in the NAML in the same period, but the difference is not very big. On the other hand, the championship concentration seems to be much higher in European football. The same is shown when it comes to qualification to post-season playoff for the Big Five leagues. Even if the differences are lesser between the minor Scandinavian leagues and NAML, this might be issues of importance in the future. One obvious difference in the qualification procedure for playoffs is the relative higher number of teams to qualify in the NAML, compared to the UEFA Champions League. However, variations among the teams to qualify for the UEFA cup might be higher than for the UEFA Champions League, and is hence reducing the potential competitive balance problem. However, this segment has not been analysed in this thesis.

The determinants for differences in competitive balance across European football leagues generally find characteristics outside the sporting league to be most significant, and not tournament structure decided by the governing body, maybe except from championship playoffs. However, there might be some indications that the distribution of revenues from sale of broadcasting rights can be relevant for the Big Five leagues. This is therefore one obvious factor that should be considered in the future. The significance of determinants for possible changes in within-league competitive balance over time differs between leagues. Even if it might be possible to affect competitive balance positively by increasing the number of teams to relegate, and also with regards to the number of teams in the top division, factors currently decided outside the league seems again to be of more importance for the domestic leagues. This means the structure of the UEFA Champions League and the labour market. However, the latter is not easy for the sport in general to affect, since it is also dependent on general political agreements. What league governing bodies might want to do is to affect decisions related to the structure of the UEFA Champions League. Here, it is important to emphasise that many changes affecting European football have happened since the middle of the 1990s, and probably we still don't know all long-term consequences yet.

For example, for the Norwegian league the same team won every championship in the first nine years in the post-Bosman period (1996-2004). However, from 2005 to 2008 four different teams have won this league. In general, competitive balance seems to have improved for all three dimensions in the Bosman II period compared to the previous period in the Norwegian top division (not significant for the SRCC). The dynasty in the first part of the post-Bosman verdict period has been partly explained by the introduction of the UEFA Champions League, and the extended revenues rewarded to one particular team. Clearly, it affected the revenue balance in the Norwegian top division. However, with the relative high probability of not qualifying for small league teams (mainly only the championship winner qualifies for the qualifying round), failure for a dominating team to qualify in a number of seasons is highly probable. One obvious long-term difficulty is hence caused by the probability of a changed cost structure, due to the extended revenue streams related to qualifications for the group play in the UEFA Champions League, because it can be difficult to reverse.[283] This might actually have a long-run positive effect on competitive balance for this type of league, related to the reduced advantage from the distribution of financial resources, if the dominating team/teams fails/fail to qualify.

The prize structure of modern professional team sports leagues might relax the statement in Cairns et al. (1986) mentioned earlier in this chapter, related to the demand for supporting a team, as might be diminishing if the team is never in the competition for the championship. This is because a team can be a winner even if it is not the champion. One example from English football is that Liverpool FC has never been the winner of the FA Premier League, but has won the UEFA Champions League. When Chelsea won the FA Premier League in 2004/05, the team became the championship winner for the first time since the 1954/55 season. In other words, the current system of European football still gives opportunity for "new" teams to win the league, and also to be winning a number of prizes during a season. However, the latter might be more relevant in bigger leagues than smaller leagues, even if smaller leagues also have different cup tournaments. Adopting a North American like playoff system (i.e. knock out tournament) might reduce the championship concentration in these leagues. Another possibility for European football leagues outside the biggest is to co-operate in a post-seasonal tournament similar to the Scandinavian Royal League. However, even if such a league could be positive for the teams' revenues and increasing the domestic leagues'

[283] See Solberg and Haugen (2009) for an example from Swedish football.

competitive intensity, it has been a failure, which is reflected in that the tournament has not been held the last two seasons.

This thesis suggests that multi-prize tournaments and the competitiveness within the tournament might have compensated eventual negative effects from bigger differences in sporting quality among the teams in domestic leagues. Even if there might still be a question about the long-run effects on competitive balance from the changes during the last 15 years, the results from the English top division suggest that relaxing the transfer system towards free agency together with the structural changes in the UEFA Champions League, have a negative effect on competitive balance, maybe except from a possible positive effect on championship concentration from the latest expansions. The latter has, on the other hand, increased the competitive intensity in for example this league. This thesis has not focused on optimal level of neither competitive balance nor competitive intensity, but there might be some trade-off point between eventually increased qualifying differences and sporting prize structure, when it comes to fan interest.

In summary, based on the time-series analysis of the English and Norwegian top divisions in football, the strength of how regulations and prize and tournament structure affect competitive balance varies between leagues and over time, with more significant reactions in the English league. Across leagues, variables outside the sport might be at least as important as sporting regulations, such as that higher GNP per capita is an explanatory factor for better win dispersion and less performance persistence. Because of the many aspects of competitive balance, broad analyses are required to give precise conclusions on a league's competitive balance level. The top division in the Norwegian league emphasises this, because when the league was least balanced in the dimension of championship concentration, both win dispersion and performance persistence dimensions ranked the league in the European top ten.

Competitive balance seems to be an inexhaustible topic. It has been subject for many analyses, but there is still a broad range of analyses left for future research. This thesis concentrates on a conceptual framework, recognising that competitive balance is a complex concept. Further, the complexity of current real-world league structure complicates the initial understanding of the competitive balance and within-league competitiveness. Future research should therefore go beyond competitive balance and focus more on the competitive intensity of professional sports leagues.

8. References

Abrevaya, J. (2004). Fit to be Tied: The Incentive Effects of Overtime Rules in Professional Hockey. *Journal of Sports Economics, 5, 292-306.*

Andreff, W. & Bourg, J. (2006). Broadcasting Rights and Competition in European Football. In C. Jeanrenaud and S. Késenne (Eds.), *The Economics of Sport and the Media.* Cheltenham, UK: Edward Elgar.

Andreff, W. & Staudohar, P. (2002). European and US sports business models. In C. Barros, M. Ibrahimo & S. Szymanski (Eds.), *Transatlantic Sport: The Comparative Economics of North American and European Sports.* Cheltenham, UK and Northampton, MA: Edward Elgar.

Atkinson, S., Stanley, L. & Tschirhart, J. (1988). Revenue Sharing as an Incentive in an Agency Problem: An Example from the National Football League. *RAND Journal of Economies, 19*, 27-43.

Baimbridge, M., Cameron, S. & Dawson, P. (1996). Satellite Television and the Demand for Football: A Whole New Ball Game? *Scottish Journal of Political Economy, 43*, 317-333.

Balfour, A. & Porter, P. (1991). The Reserve Clause in Professional Sports. Legality and Effect on Competitive Balance. *Labor Law Journal, 42*, 8-18.

Barros, C., Ibrahimo, M. & Szymanski, S. (2002). Transatlantic Sport: An Introduction. In C. Barros, M. Ibrahimo & S. Szymanski (Eds.), *Transatlantic Sport: The Comparative Economics of North American and European Sports.* Cheltenham, UK and Northampton, MA: Edward Elgar.

Bennett, R. & Fizel, J. (1995). Telecast Deregulation and Competitive Balance: Regarding NCAA Division I Football. *American Journal of Economics and Sociology, 54*, 183-199.

Berri, D., Brook, S., Frick, B., Fenn, A. & Vicente-Mayoral, R. (2005). The Short Supply of Tall People: Competitive Imbalance and the National Basketball Association. *Journal of Economic Issues, XXXIX*, 1029-41.

Bird, P. (1982). The Demand for League Football. *Applied Economics, 14*, 637-649.

Borland, J. (1987). The Demand for Australian Rules Football. *The Economic Record, 63*, 221-230.

Borland, J. & Lye, J. (1992). Attendance at Australian Rules Football: A Panel Study. *Applied Economics, 24*, 1053-1058.

Borland, J. & Macdonald, R. (2003). Demand for Sport. *Oxford Review of Economic Policy, 19*, 478-502.

Bourgheas, S. & Downward, P. (2003). The Economics of Professional Sports Leagues: Some Insights on the Reform of Transfer Markets. *Journal of Sports Economics, 4*, 87-107.

Bourg, J. (2004). Professional Team Sports in Europe: Which Economic Model? In R. Fort and J. Fizel (Eds.), *International Sports Economics Comparisons.* Westport, CT: Praeger Publishing.

Brocas, I. & Carrillo, J. (2004). Do the "Three-Point Victory" and "Golden Goal" Rules Make Soccer More Exiting? *Journal of Sports Economics, 5*, 169-185.

Brown, T. (1995). *A Short History of the Football League.* Retrieved by April 2008, from www.soccer.mistral.co.uk/fleg/flhist.htm.

Buraimo, B., Simmons, R. & Szymanski, S. (2006). English Football. *Journal of Sports Economics, 7*, 29-46.

Butler, M. (1995). Competitive Balance in Major League Baseball. *The American Economist, 39*, 46-52.

Buzzacchi, L., Szymanski, S. & Valletti, T. (2003). Equality of Opportunity and Equality of Outcome: Open Leagues, Closed Leagues and Competitive Balance. *Journal of Industry, Competition and Trade, 3*, 167-186.

Cain, L. & Haddock, D. (2006). Measuring Parity: Tying Into the Idealized Standard Deviation. *Journal of Sports Economics, 7*, 330-338.

Cairns, J. (1987). Evaluating Changes in League Structure: The Reorganization of the Scottish Football League. *Applied Economics, 19*, 259-275.

Cairns, J. (1988). Uncertainty of Outcome and the Demand for Football. Discussion Paper No. 88-02. University of Aberdeen.

Cairns, J. (1990). The Demand for Professional Team Sports. *British Review of Economic Issues, 12*, 1-20.

Cairns, J., Jennett, N. & Sloane, P. (1986). The Economics of Professional Team Sports: A Survey of Theory and Evidence. *Journal of Economic Studies, 13*, 3-80.

Caves, R. & Porter, M. (1977). From Entry Barriers to Mobility Barriers: Conjectural Decisions and Contrived Deterrence to New Competition. *Quarterly Journal of Economics, 91*, 241-261.

Coase, R. (1960). The Problem of Social Cost. *Journal of Law and Economics, 3*, 1-44.

Copeland, T., Weston, J. & Shastri, K. (2005). *Financial Theory and Corporate Theory*. Boston, MA: Pearson Addison-Wesley.

Cyrenne, P. (2001). A Quality-of-Play Model of a Professional Sports League. *Economic Inquiry, 39*, 444-452.

Czarnitzki, D. & Stadtmann, G. (2002). Uncertainty of Outcome versus Reputation: Empirical Evidence for the First German Football Division. *Empirical Economics, 27*, 101-112.

Dabscheck, B. (1975a). Sporting Equality: Labour Market versus Product Market Control. *Journal of Industrial Relations, 17*, 174-190.

Dabscheck, B. (1975b). The Wage Determination Process for Sportsmen. *Economic Record, 51*, 52-65.

Daly, G. & Moore, W. (1981). Externalities, Property Rights and the Allocation of Resources in Major League Baseball. *Economic Inquiry, 19*, 77-95.

Davenport, D. (1969). Collusive Competition in Major League Baseball. Its Theory and Institutional Development. *American Economist, 13*, 6-30.

Davis, L. (1974). Self-Regulation in Baseball, 1909-71. In R. Noll (Ed.), *Government and the Sports Business*. Washington, DC: Brookings Institution.

Dawson, A. & Downward, P. (2005). Measuring Short-Run Uncertainty of Outcome in Sporting Leagues: A Comment. *Journal of Sports Economics, 5*, 303-313.

Dejonghe, T. (2004). *Sport en Economie: Een Noodzaak tot Symbiose"*. Nieuwegein: Arko Sports Media.

Deloitte & Touche (2000). Annual Review of Football Finance: A Review of the 1998/1999 Season. Manchester: Deloitte and Touche Sport.

Demmert, H. (1973). *The Economics of Professional Team Sports*. Lexington, MA: Lexington Books.

Depken, C. (1999). Free-Agency and the Competitiveness of Major League Baseball. *Review of Industrial Organization, 14*, 205-217.

Depken, C. (2002). Free Agency and the Concentration of Player Talent in Major League Baseball. *Journal of Sports Economics, 3*, 335-353.

Dobson, S. & Gerrard, B. (1999). The Determination of Player Transfer Fees in English Professional Soccer. *Journal of Sport Management, 13*, 259-279.

Dobson, S. & Gerrard, B. (2000). Testing for Monopoly Rents in the Market for Playing Talent: Evidence from English Professional Football. *Journal of Economic Studies, 27*, 142-164.

Dobson, S. & Goddard, J. (1992). The Demand for Standing and Seated Viewing Accommodation in the English Football League. *Applied Economics, 24*, 1155-1163.

Dobson, S. & Goddard, J. (1995). The Demand for Professional League Football in England and Wales, 1925-92. *The Statistician, 44*, 259-277.

Dobson, S. & Goddard, J. (1996). The Demand for Football in the Regions of England and Wales. *Regional Studies, 30*, 443-453.

Dobson, S. & Goddard, J. (1998). Performance, Revenue, and Cross subsidization in the Football League, 1927-1994. *Economic History Review, LI*, 763-785.

Dobson, S. & Goddard, J. (2001). *The Economics of Football*. Cambridge, UK: Cambridge University Press.

Dobson, S. & Goddard, J. (2004). Revenue Divergence and Competitive Balance in a Divisional Sports League. *Scottish Journal of Political Economy, 51*, 359-376.

Doornik, J. & Hendry, D. (2006a). *Empirical Econometric Modelling. PcGiveTM11: Volume I*. London, UK: Timberlake Consultants Ltd.

Doornik, J. & Hendry, D. (2006b). *Modelling Dynamic Systems: PcGiveTM11: Volume III*. London, UK: Timberlake Consultants Ltd.

Downward, P. & Dawson, A. (2000). *The Economics of Professional Team Sports*. London, UK and New York, NY: Routledge.

Easton, S. & Rockerbie, D. (2005). Overtime! Rules and Incentives in the National Hockey League. *Journal of Sports Economics, 6*, 178-202.

Eckard, E. (1998). The NCAA Cartel and Competitive Balance in College Football. *Review of Industrial Organization, 13*, 347-369.

Eckard, E. (2001). Free Agency, Competitive Balance and Diminishing Returns to Pennant Contention. *Economic Inquiry, 39*, 430-443.

Eckard, E. (2003). The ANOVA-Based Competitive Balance Measure. *Journal of Sports Economics, 4*, 74-80.

Eggen, N. & Nyrønning, S. (1999). *Godfoten: Samhandling - Veien til Suksess*. Oslo, Norway: Aschehoug.

El-Hodiri, M. & Quirk, J. (1971). An Economic Model of a Professional Sports League. *Journal of Political Economy, 79*, 1302-1309.

Enders, W. (2004). *Applied Econometric Time Series*. Second Edition. Hoboken, NJ: Wiley.

Falter, J. & Perignon, C. (2000). Demand for Football and Intramatch Winning Probability: An Essay on the Glorious Uncertainty of Sports. *Applied Economics, 32*, 1757-1765.

Feigenbaum, A. & Primeaux, W. (1983). Strategic Groups and Mobility Barriers: The Level of Struggle in an Industry. Working Paper No. 984, University of Illinois.

Fernandez-Cantelli, E. & Meeden, G. (2002). An Improved Award System for Soccer. Retrieved by May 2008, from www.stat.umn.edu/~glen/papers/soccer5.pdf.

Forrest, D., Beaumont, J., Goddard, J. & Simmons, R. (2005). Home Advantage and the Debate About Competitive Balance in Professional Sports Leagues. *Journal of Sports Science, 23*, 439-445.

Forrest, D. & Simmons, R. (2002). Outcome Uncertainty and Attendance Demand in Sport: The Case of English Soccer. *The Statistician, 51*, 229-241.

Forrest, D., Simmons, R. & Buraimo, B. (2006). Broadcaster and Audience Demand for Premier League Football. In C. Jeanrenaud & S. Késenne (Eds.), *The Economics of Sport and the Media*. Cheltenham, UK: Edward Elgar.

Fort, R. (2000). European and North American Sports Differences (?). *Scottish Journal of Political Economy, 47*, 431-455.

Fort, R. (2003). *Sports Economics*. Upper Saddle River, NJ: Prentice-Hall.

Fort, R. (2006a). *Sports Economics*. Second Edition. Upper Saddle River, NJ: Prentice-Hall.

Fort, R. (2006b). Talent Market Models in North American and World Leagues. In P. Rodrígues, S. Késenne & J. García (Eds.), *Sports Economics after Fifty Years: Essays in Honour of Simon Rottenberg*. Oviedo. Spain: Oviedo University Press.

Fort, R. (2007). Comments on "Measuring Parity". *Journal of Sports Economics, 8*, 642-651.

Fort, R. & Fizel, J. (2004). Preface. In R. Fort & J. Fizel (Eds.), *International Sports Economics Comparisons*. Westport, CT. Praeger Publishing.

Fort, R. & Lee, Y. (2006). Stationarity and Major League Baseball Attendance Analysis. *Journal of Sports Economics, 7*, 408-415.

Fort, R. & Lee, Y. (2007). Structural Change, Competitive Balance, and the Rest of the Major Leagues. *Economic Inquiry, 45*, 519-532.

Fort, R. & Maxcy, J. (2003). Comment: Competitive Balance in Sports Leagues: An Introduction. *Journal of Sports Economics, 4*, 154-160.

Fort, R. & Quirk, J. (1995). Cross-subsidization, Incentives, and Outcomes in Professional Team Sports Leagues. *Journal of Economic Literature, XXXIII*, 1265-1299.

Fort, R. & Quirk, J. (1997). Introducing a Competitive Economic Environment into Professional Sports. In W. Hendricks (Ed.), *Advances in the Economics of Sports, Volume 2*. Greenwich, CT: JAI Press.

Frank, R. & Cook, P. (1996). *The Winner-Take-All Society*. New York: Penguin.

Frick, B. & Prinz, J. (2004). Revenue-Sharing Arrangements and the Survival of Promoted Teams: Empirical Evidence from the Major European Soccer Leagues. In R. Fort & J. Fizel (Eds.), *International Sports Economics Comparisons*. Westport, CT: Praeger Publishing.

Gammelsæter, H. & Ohr, F. (2002). *Kampen uten Ball. Om Penger, Ledelse og Identitet i Norsk Fotball*. Oslo, Norway: Abstrakt Forlag.

Garcia, J. & Rodriguez, P. (2002). The Determinants of Football Match Attendance Revisited: Empirical Evidence From the Spanish Football League. *Journal of Sports Economies, 3*, 18-38.

Gaviria, A. (2000). Is Soccer Dying? A Time Series Approach. *Applied Economics Letters, 7*, 275-278.

Gerrard, B. (1998). Competitive Balance in the F.A. Premier League, unpublished mimeo, Leeds University Business School.

Gerrard, B. (1999). Team Sports as a Free-market Commodity. *New Political Economy, 4*, 273-278.

Gerrard, B. (2000a). Lecture notes about "Nonstationary Time Series". Leeds University Business School.

Gerrard, B. (2000b). Lecture notes about "Player Wages" in the course BUEC2450 "The economics and finance of Football". Leeds University Business School.

Gerrard, B. (2000c). Media Ownership of Pro Sports Teams: Who are the Winners and Losers? *International Journal of Sports Marketing and Sponsorship, 2*, 199-218.

Gerrard, B. (2001). A New Approach to Measuring Player and Team Quality in Professional Team Sports. *European Sport Management Quarterly, 1*, 219-234.

Gerrard, B. (2002). The Muscle Drain, Coubertobin-Type Taxes and the International Transfer System in Association Football. *European Sport Management Quarterly, 2*, 47-56.

Gerrard, B. (2004a). The Measurement and Valuation of Player Quality in Association Football, unpublished mimeo, Leeds University Business School.

Gerrard, B. (2004b). Still up for Grabs? Maintaining the Sporting and Financial Viability of European Club Soccer. In R. Fort & J. Fizel (Eds.), *International Sports Economics Comparisons*. Westport, CT: Praeger Publishing.

Gerrard, B. (2005). A Resource-Utilization Model of Organizational Efficiency in Professional Sports Teams. *Journal of Sport Management, 19*, 143-169.

Gerrard, B. (2006a). Analysing the Win-Wage Relationship in Pro Sports Leagues: Evidence from the FA Premier League, 1997/98 – 2001/02. In P. Rodrígues, S. Késenne & J. García (Eds.), *Sports Economics after Fifty Years: Essays in Honour of Simon Rottenberg*. Oviedo, Spain: Oviedo University Press.

Gerrard, B. (2006b). Competitive Balance and the Sports Media Rights Market: What are the Real Issues? In C. Jeanrenaud & S. Késenne (Eds), *The Economics of Sport and the Media*. Cheltenham, UK: Edward Elgar.

Goksøyr, M. & Olstad, F. (2002). *Fotball! Norges Fotballforbund 100 år*. Oslo, Norway: Norges Fotballforbund.

Goossens, K. (2006). Descriptive Analysis of Competitive Balance Within and Between European Football Leagues. IASE Conference, Bochum, Germany.

Goossens, K. & Késenne, S. (2007). National Dominance in European Football Leagues. In M. Parent & T. Slack (Eds.), *International Perspectives on the Management of Sport*. Burlington, MA, San Diego, CA and London, UK. Elsevier.

Gould, S. (1986, August). Entropic Homogeneity Isn't Why No One Hits 0.400 Any More. Discover, 60-66.

Gould, S. (1996). *Full House: The Spread of Excellence from Plato to Darwin*. New York, NY: Three Rivers Press.

Gratton, C. & Solberg, H. (2004). Sports and Broadcasting: Comparisons between the United States and Europe. In R. Fort & J. Fizel (Eds.), *International Sports Economics Comparisons*. Westport, CT: Praeger Publishing.

Gratton, C. & Solberg, H. (2007). *The Economics of Sports Broadcasting*. London, UK and New York, NY: Routledge.

Green, G. (1953). *The History of the Football Association*. London, UK: Naldrett Press for the Football Association.

Greene, W. (1993). *Econometric Analysis*. Second Edition. New York, NY: Macmillian.

Greene, W. (2008). *Econometric Analysis*. Sixth Edition. Upper Saddle River, NJ: Prentice Hall.

Grier, K. & Tollison, R. (1994). The Rookie Draft and Competitive Balance: The Case of Professional Football. *Journal of Economic Behavior and Organization, 25*, 293-298.

Groot, L. (2008). *Economics, Uncertainty and European Football. Trends in Competitive Balance*. Cheltenham, UK and Northampton, MA: Edward Elgar.

Gärtner, M. & Pommerehne, W. (1978). Der Fussballzuschauer - Ein Homo Oeconomicus? Eine Theoretische und Empirische Analyse. *Jahrbuch für Sozialwissenschaften, 29*, 1089-1095.

Hadley, L., Ciecka, J. & Krautmann, A. (2005). Competitive Balance in the Aftermath of the 1994 Players' Strike. *Journal of Sports Economics, 6*, 379-389.

Hall, S., Szymanski, S. & Zimbalist, A. (2002). Testing Causality between Team Performance and Payroll: The Cases of Major League Baseball and English Soccer. *Journal of Sports Economics, 3*, 149-168.

Hart, R., Hutton, J. & Sharot, T. (1975). A Statistical Analysis of Association Football Attendances. *Applied Statistics, 24*, 17-27.

Haugen, K. (2008). Point Score Systems and Competitive Imbalance in Professional Soccer. *Journal of Sports Economics, 9*, 191-210.

Hausman, J. & Leonard, G. (1997). Superstars in the National Basketball Association: Economic Value and Policy. *Journal of Labor Economics, 15*, 586-624.

Hendry, D. & Juselius, K. (2000). Explaining Cointegration Analysis: Part 1. *The Energy Journal, 21*, 1-42.

Hjelseth, A. (2006). Mellom Børs, Katedral og Karneval: Norske Supporteres Forhandlinger om Kommersialisering av Fotball. Dr.polit-avhandling (PhD Dissertation). Det samfunnsvitenskapelige fakultet. Universitetet i Bergen, Norway.

Hoehn, T. & Szymanski, S. (1999). The Americanization of European Football. *Economic Policy, 28*, 205-240.

Horowitz, I. (1997). The Increasing Competitive Balance in Major League Baseball. *Review of Industrial Organization, 12*, 373-387.

Humphreys, B. (2002). Alternative Measures of Competitive Balance in Sports Leagues. *Journal of Sports Economics, 3*, 133-148.

Humphreys, B. (2003). Comments on "Thinking about Competitive Balance". *Journal of Sports Economics, 4*, 284-287.

Janssens, P. & Késenne, S. (1987). Belgian Soccer Attendances. *Tjidschrift voor Economie en Management, XXXII*, 305-315.

Jeanrenaud, C. & Késenne, S. (2006). Sport and the Media: An Overview. In C. Jeanrenaud & S. Késenne (Eds.), *The Economics of Sport and the Media*. Cheltenham, UK: Edward Elgar.

Jennett, N. (1984). Attendances, Uncertainty of Outcome and Policy in Scottish League Football. *Scottish Journal of Political Economy, 31*, 176-198.

Jones, J. (1969). The Economics of the National Hockey League. *Canadian Journal of Economics, 2*, 1-20.

Jones, J. & Ferguson, D. (1988). Location and Survival in the National Hockey League. *The Journal of Industrial Economics, XXXVI*, 443-457.

Kahane, L. (2003). Comments on "Thinking about Competitive Balance". *Journal of Sports Economics, 4*, 288-291.

Kamerschen, D. & Lam, N. (1975). A Survey of Measures of Market Power. *Rivista Internazionale di Scienze Economiche e Commerciali, 22*, 1131-1156.

Késenne, S. (1996). League Management in Professional Team Sports with Win Maximizing Clubs. *European Journal for Sports Management, 2*, 14-22.

Késenne, S. (1999). Player Market Regulation and Competitive Balance in a Win Maximizing Scenario. In C. Jeanrenaud & S. Késenne (Eds.), *Competition Policy in Professional Sports: Europe after the Bosman Case*. Antwerp, Belgium: CIES, Standaard Editions.

Késenne, S. (2000a). The Impact of Salary Caps in Professional Team Sports. *Scottish Journal of Political Economy, 47*, 422-430.

Késenne, S. (2000b). Revenue Sharing and Competitive Balance in Professional Team Sports. *Journal of Sports Economics, 1*, 56-65.

Késenne, S. (2002a). Improving the Competitive Balance and the Salary Distribution in Professional Team Sports. In C. Barros, M. Ibrahimo & S. Szymanski (Eds.), *Transatlantic Sport: The Comparative Economics of North American and European Sports*. Cheltenham, UK and Northampton, MA: Edward Elgar.

Késenne, S. (2002b). The Monopsonistic Player Labour Market in a Win-Maximising League. *European Sport Management Quarterly, 2*, 180-187.

Késenne, S. (2004). Competitive Balance and Revenue Sharing: When Rich Clubs Have Poor Teams. *Journal of Sports Economics, 5*, 206-212.

Késenne, S. (2005a). Revenue Sharing and Competitive Balance. Does the Invariance Proposition Hold? *Journal of Sports Economics, 6*, 98-106.

Késenne, S. (2005b).The Organisation of European Football and the Competitive Balance within and between Nations. Dept. of Economics, University of Antwerp, Belgium.

Késenne, S. (2006a). Competitive Balance in Team Sports and the Impact of Revenue Sharing. *Journal of Sport Management, 20*, 39-51.

Késenne, S. (2006b). The Win Maximization Model Reconsidered: Flexible Talent Supply and Efficiency Wages. *Journal of Sports Economics, 7*, 416-427.

Késenne, S. (2007a). Revenue Sharing and Owner Profits in Professional Team Sports. *Journal of Sports Economics, 8*, 519-529.

Késenne, S. (2007b). *The Economic Theory of Professional Team Sports. An Analytical Treatment.* Cheltenham, UK and Northampton, MA: Edward Elgar.

Késenne, S. (2007c). The Peculiar International Economics of Professional Team Sports. *Scottish Journal of Political Economy, 54*, 388-399.

Kipker, I. (2000). Determinanten der Zuschauernachfrage im Professionellen Teamsport: Wie Wichtig ist die Sportliche Ausgeglichenheit?, PhD Dissertiation. Henrich Heine University, Dusseldorf.

Knowles, G., Sherony, K. & Haupert, M. (1992). The demand for Major League Baseball: A Test of the Uncertainty of Outcome Hypothesis. *The American Economist, 36*, 72-80.

Koning, R. (2000). Balance in Competition in Dutch Soccer. *The Statistician, 49*, 419-431.

Kuypers, T. (1997).The Beautiful Game? An Econometric Study of Audiences, Gambling and Efficiency in English Football. PhD Dissertation. University College, London.

La Croix, S. & Kawaura, A. (1999). Rule Changes and Competitive Balance in Japanese Professional Baseball. *Economic Inquiry, 37*, 353-368.

Larsen, A., Fenn, A. & Spenner, E. (2006). The Impact of Free Agency and the Salary Cap on Competitive Balance in the National Football League. *Journal of Sports Economics, 7*, 374-390.

Lee, Y. (2004). Competitive Balance and Attendance in Japanese, Korean, and U.S. Professional Baseball Leagues. In R. Fort & J. Fizel (Eds.), *International Sports Economies Comparisons*. Westport, CT: Praeger Publishers.

Lee, Y. & Fort, R. (2005). Structural Change in MLB Competitive Balance: The Depression, Team Location, and Integration. *Economic Inquiry, 43*, 158-169.

Leeds, M. & von Allmen, P. (2002). *The Economics of Sport.* Boston, MA. Pearson Education.

Lenten, L. (2007). The Extent to Which Unbalanced Schedules Affect Reported Competitive Balance in the Scottish Premier League. Department of Economics and Finance, La Trobe University. Retrieved by April 2008, from econ.core.hu/doc/seminar/Liam_Lenten_ScottishPremierLeague.pdf.

Levin, R., Mitchell, G., Volcker, P. & Will, G. (2000). *The Report of the Independent Members of the Commissioner's Blue Ribbon Panel on Baseball Economics.* Major League Baseball. Retrieved by April 2008, from www.mlb.com/mlb/downloads/blue_ribbon.pdf

Lewis, H., Sexton, T. & Lock, K. (2007). Player Salaries, Organizational Efficiency, and Competitiveness in Major League Baseball. *Journal of Sports Economics, 8*, 266-294.

Macdonald, R. & Borland, J. (2004). Professional Sports Competitions in Australia. In R. Fort & J. Fizel (Eds.), *International Sports Economics Comparisons.* Westport, CT: Praeger Publishing.

Magee, J. & Sugden, J. (2002). "The World at their Feet": Professional Football and International Labor Migration. *Journal of Sport and Social Issues, 26*, 421-437.

Maguire, J. (1996). Blade Runners: Canadian Migrants, Ice Hockey and the Global Sports Process. *Journal of Sport and Social Issues, 20*, 335-360.

Maguire, J. (1999). *Global Sport: Identities, Societies, Civilizations*. Cambridge: Polity.

Marburger, D. (1997). Gate Revenue Sharing and Luxury Taxes in Professional Sports. *Contemporary Economic Policy, XV*, 114-123.

Marburger, D. (2002). Property Rights and Unilateral Player Transfers in a Multiconference Sports League. *Journal of Sports Economics, 3*, 122-132.

Mason, C. (1997). Revenue Sharing and Agency Problems in Professional Team Sport: The Case of the National Football League. *Journal of Sport Management, 11*, 203-222.

Massarella, L. (2000, February). Foreign Disgrace. FourFourTwo, 62-65.

Maxcy, J. (2002). Rethinking Restrictions on Player Mobility in Major League Baseball. *Contemporary Economic Policy, 20*, 145-159.

Maxcy, J. & Mondello, M. (2006). The Impact of Free Agency on Competitive Balance in North American Professional Team Sports Leagues. *Journal of Sport Management, 20*, 345-365.

McArdle, D. (2000). One Hundred Years of Servitude: Contractual Conflict in English Professional Football before *Bosman*. *Web Journal of Current Legal Issues*. Retrieved by April 2008, from webjcli.ncl.ac.uk/2000/issue2/mcardle2.html.

McCoy, J. (Ed.). (2002). *Geo-Data: The World Geographical Encyclopaedia*. Third Edition. Detroit, MI: Thomson-Gale.

Michie, J. & Oughton, C. (2004). *Competitive Balance in Football: Trends and Effects*. London, UK: The Sports Nexus.

Morgan, S. (1999, May). National Disgrace. FourFourTwo, 67-76.

Morrow, S. (1999). *The New Business of Football, Accountability and Finance in Football*. London. Macmillan.

Nawrat, C. & Hutchings, S. (1998). *The Sunday Times Illustrated History of Football*. London, UK: Hamlyn.

Neale, W. (1964). The Peculiar Economics of Professional Sports. *The Quarterly Journal of Economics, LXXVIII*, 1-14.

Newson, G. (1984). Three Points for a Win: Has It Made Any Difference? *The Mathematical Gazette, 68*, 87-91.

Noether, G. (1981). Why Kendall Tau? *Teaching Statistics, 3*, 41-43.

Noll, R. (1974). Attendance and Price Setting. In R. Noll (Ed.), *Government and the Sports Business*. Washington, DC: Brookings.

Noll, R. (1982). Major League Team Sports. In W. Adams (Ed.), *The Structure of American Industry*. Sixth edition. New York, NY and London, UK: Macmillian.

Noll, R. (1988). *Professional basketball* (Studies in Industrial Economics Paper No 144). Stanford University, Stanford, California.

Noll, R. (2002). The Economics of Promotion and Relegation in Sports Leagues: The Case of English Football. *Journal of Sports Economics, 3*, 169-203.

Noll, R. (2003). The Organization of Sports Leagues. *Oxford Review of Economic Policy, 19*, 530-551.

Noll, R. (2006). Sports Economics after Fifty Years. In P. Rodrígues, S. Késenne & J. García (Eds.), *Sports Economics after Fifty Years: Essays in Honour of Simon Rottenberg*. Oviedo, Spain: Oviedo University Press.

Ohr, F. & Solum, N. (2001). Tippeligaen 2000: Økonomiske Nøkkeltall. Arbeidsrapport 0112. Møreforsking Molde, Norway.

Owen, P. & Weatherston, C. (2002). Professionalization of New Zealand Rugby Union: Historical Background, Structural Changes and Competitive Balance. Economics Discussion Papers No. 0214: University of Otago, New Zealand.

Owen, P. & Weatherston, C. (2004). Professionalization and Competitive Balance in New Zealand Rugby Union. In R. Fort & J. Fizel (Eds.), *International Sports Economics Comparisons*. Westport, CT: Praeger Publishing.
Palacios-Huerta, I. (2004). Structural Changes During a Century of the World's Most Popular Sport. *Statistical Methods and Applications, 13*, 241-258.
Paul, R. (2003). Variations in NHL Attendance: The Impact of Violence, Scoring and Regional Rivalries. *The American Journal of Economics and Sociology, 62*, 345-364.
Paul, R., Weinbach, A. & Melvin, P. (2004). The Yankees Effect: The Impact of Interleague Play and the Unbalanced Schedule on Major League Baseball Attendance. *New York Economic Review, 35*, 3-15.
Peel, D. & Thomas, D. (1988). Outcome uncertainty and the Demand for Football: An Analysis of Match Attendances in the English Football League. *Scottish Journal of Political Economy, 35*, 242-249.
Peel, D. & Thomas, D. (1992). The Demand for Football: Some Evidence on Outcome Uncertainty. *Empirical Economics, 17*, 323-331.
Peel, D. & Thomas, D. (1996). Attendance Demand: An Investigation of Repeat Fixtures. *Applied Economics Letters, 3*, 391-394.
Peel, D. & Thomas, D. (1997). Handicaps, Outcome Uncertainty and Attendance Demand. *Applied Economics Letters, 4*, 567-570.
Porter, P. (1992). The Role of the Fan in Professional Baseball. In P. Sommers (Ed.), *Diamonds are Forever: The Business of Baseball*. Washington, DC: Brookings.
Quirk, J. & El-Hodiri, M. (1974). The Economic Theory of a Professional Sports League. In R. Noll (Ed), *Government and the Sports Business*. Washington, DC: Brookings Institution.
Quirk, J. & Fort, R. (1992). *Pay Dirt. The Business of Professional Team Sports*. Princeton, NJ: Princeton University Press.
Rascher, D. (1997). A Model of Professional Sports League. In W. Hendricks (Ed.), *Advances in the Economics of Sport (Vol. 2)*. Greenwich, CT: JAI Press.
Ross, S. & Lucke, R. (1997). Why Highly Paid Athletes Deserve More Antitrust Protection than Ordinary Unionized Workers. *Antitrust Bulletin, 42*, 641-679.
Rottenberg, S. (1956). The Baseball Players' Labor Market. *Journal of Political Economy, 64*, 242-258.
Ryder, A. (2004a). Poisson Toolbox: A Review of the Application of the Poisson Probability Distribution in Hockey; www.HockeyAnalytics.com.
Ryder, A. (2004b). Win Probabilities; A Tour Trough Win Probability Models for Hockey; www.HockeyAnalytics.com.
Sanderson, A. (2002). The Many Dimensions of Competitive Balance. *Journal of Sports Economics, 3*, 204-228.
Sanderson, A. & Siegfried, J. (2003). Thinking About Competitive Balance. *Journal of Sports Economics, 4*, 255-279.
Sandy, R., Sloane, P. & Rosentraub, M. (2004). *The Economics of Sport: An International Perspective*. New York, NY: Palgrave MacMillan.
Schmidt, M. (2001). Competition in Major League Baseball: The Impact of Expansion. *Applied Economics Letters, 8*, 21-26.
Schmidt, M. & Berri, D. (2001). Competitive Balance and Attendance: The Case of Major League Baseball. *Journal of Sports Economics, 2*, 145-167.
Schmidt, M. & Berri, D. (2002). The Impact of the 1981 and 1994-1995 Strikes on Major League Baseball Attendance: A Time-Series Analysis. *Applied Economics Letters, 34*, 471-478.
Schmidt, M. & Berri, D. (2003). On the Evolution of Competitive Balance: The Impact of an Increasing Global Search. *Economic Inquiry, 41*, 692-704.

Schmidt, M. & Berri, D. (2005). Concentration of Playing Talent: Evolution in Major League Baseball. *Journal of Sports Economics, 6*, 412-419.

Scully, G. (1974). Pay and Performance in Major League Football. *The American Economic Review, 64*, 915-930.

Scully, G. (1989). *The Business of Major League Baseball*. Chicago, IL: University of Chicago Press.

Scully, G. (1995). *The Market Structure of Sports*, Chicago, IL: University of Chicago Press.

Scully, G. (2002). The Distribution of Performance and Earnings in a Prize Economy. *Journal of Sports Economics, 3*, 235-245.

Scully, G. (2004). The Market Structure of Sports. In S. Rosner & K. Shropshire (Eds.), *The Business of Sports*. Sudbury, MA: Jones and Barlett Publishers.

Serieforeningen av 1972 (1997). Skistad, B. (Ed.), *Serieforeningen av 1972 25 år*. Drammen, Norway.

Shank, M. (2002). *Sports Marketing: A Strategic Perspective*. Second Edition. Upper Saddle River, NJ: Prentice Hall.

Shmanske, S. & Lowenthal, F. (2007). Overtime Incentives in the National Hockey League (NHL): More Evidence. *Journal of Sports Economics, 8*, 435-442.

Simmons, R. (1996). The demand for English League Football: A Club-Level Analysis. *Applied Economics, 28*, 139-155.

Sloane, P. (1969). The Labour Marked in Professional Football. *British Journal of Industrial Relations, 7*, 181-199.

Sloane, P. (1971). The Economics of Professional Football: The Football Club as a Utility Maximiser. *Scottish Journal of Political Economy, 18*, 121-146.

Sloane, P. (1976a). Restriction on Competition in Professional Team Sports. *Bulletin of Economic Research, 28*, 3-22.

Sloane, P. (1976b). Sporting Equality: Labour Market versus Product Market Control - A Comment. *Journal of Industrial Relations, 18*, 79-84.

Sloane, P. (1980). *Sport in the Market?* Hobart Paper No. 85. London: Institute of Economic Affairs.

Sloane, P. (2002). The Regulation of Professional Team Sports. In C. Barros, M. Ibrahimo & S. Szymanski (Eds.), *Transatlantic Sport; The Comparative Economics of North American and European Sports*. Cheltenham, UK and Northampton, MA. Edward Elgar.

Sloane, P. (2006). Rottenberg and the Economics of Sport after 50 Years: An Evaluation. In P. Rodrígues, S. Késenne & J. García (Eds.), *Sports Economics after Fifty Years: Essays in Honour of Simon Rottenberg*. Oviedo, Spain: Oviedo University Press.

Smart, R. & Goddard, J. (1991). The Determinants of Standing and Seated Football Attendances: Evidence from Three Scottish League Clubs. *Quarterly Economic Commentary, 16*, 61-64.

Solberg, H. (2004). Sport Broadcasting. In J. Beech & S. Chadwick (Eds.), *The Business of Sport Management*. London, UK: Prentice Hall.

Solberg, H. & Haugen, K. (2009). European Club Football - Why Enormous Revenues are not enough? *Sport in Society*, forthcoming.

Staudohar, P. (1999). Salary Caps in Professional Team Sports. In C. Jeanrenaud & S. Késenne (Eds.), *Competition Policy in Professional Sports: Europe after the Bosman Case*. Antwerp, Belgium: CIES, Standaard Editions.

Stefani, R. (1983). Observed Betting Tendencies and Suggested Betting Strategies for European Football Pools. *The Statistician, 32*, 319-329.

Sutter, D. & Winkler, S. (2003). NCAA Scholarship Limits and Competitive Balance in College Football. *Journal of Sports Economics, 4*, 3-18.

Sutton, W., McDonald, M., Milne, G. & Cimperman, J. (1997). Creating and Fostering Fan Identification in Professional Sports. *Sport Marketing Quarterly, 6*, 15-22.

Szymanski, S. (1998). Hermetic Leagues, Open Leagues and the Impact of Revenue Sharing on Competitive Balance. Paper submitted as evidence to the Restrictive Practices Court in the Premier League/Sky/BBC case.

Szymanski, S. (2001). Income Inequality, Competitive Balance and the Attractiveness of Team Sports: Some Evidence and a Natural Experiment from English Soccer. *The Economic Journal, 111*, F69-F84.

Szymanski, S. (2002). The Invariance Principle, Prizes and Competitive Balance. In G. Papanikos (Ed.), *The Economics and Management of Mega Athletic Events: Olympic Games, Professional Sports and Other Essays*. Athens, Greece: ATINER.

Szymanski, S. (2003a). The Assessment: The Economics of Sport. *Oxford Review of Economic Policy, 19*, 467-477.

Szymanski, S. (2003b). The Economic Design of Sporting Contests. *Journal of Economic Literature, XLI*, 1137-1187.

Szymanski, S. (2003c). Incentives and Competitive Balance in Team Sports. *European Sport Management Quarterly, 3*, 11-30.

Szymanski, S. (2004). Is There a European Model of Sports? In R. Fort & J. Fizel (Eds.), *International Sports Economics Comparisons*. Westport, CT: Praeger Publishing.

Szymanski, S. & Késenne, S. (2004). Competitive Balance and Gate Revenue Sharing in Team Sports. *The Journal of Industrial Economics, LII, 165-177*.

Szymanski, S. & Kuypers, T. (1999). *Winners and Losers: The Business Strategy of Football*. London, UK: Viking Books.

Szymanski, S. & Ross, S. (2000). Open Competition in League Sports. U Illinois Law & Economics Research Paper No. 00-07. Retrieved by April 2008, from SSRN: http://ssrn.com/abstract=243756.

Szymanski, S. & Smith, R. (1997). The English Football Industry: Profit, Performance and Industrial Structure. *International Review of Applied Economics, 11*, 135-153.

Szymanski, S. & Smith, R. (2002). Equality of Opportunity and Equality of Outcome: Static and Dynamic Competitive Balance in European and American Sports Leagues. In C. Barros, M. Ibrahimio & S. Szymanski (Eds.), *Transatlantic Sport: The Comparative Economics of North American and European Sports*. Cheltenham, UK: Edward Elgar.

Szymanski, S. & Valletti, T. (2005a). Incentive Effects of Second Prizes. *European Journal of Political Economy, 21*, 467-481.

Szymanski, S. & Valletti, T. (2005b). Promotion and Relegation in Sporting Contests. *Rivista di Politica Economica, 95*, 3-39.

Szymanski, S. & Zimbalist, A. (2005). *National Pastime: How Americans Play Baseball and the Rest of the World Plays Soccer*. Washington, DC: Brookings Institution Press.

Taylor, B. & Trogdon, J. (2002). Losing to Win: Tournament Incentives in the National Basketball Association. *Journal of Labor Economics, 20*, 23-41.

Tonazzi, A. (2003). Competition Policy and the Commercialization of Sport Broadcasting Rights: The Decision of the Italian Competition Authority. *International Journal of the Economics and Business, 10*, 17-34.

Torgler, B. (2007). "La Grande Boucle": Determinants of Success at the Tour de France. *Journal of Sports Economics, 8*, 317-331.

Triola, M. & Franklin, L. (1994). *Business Statistics: Understanding Populations and Processes*. Reading, MA: Addison-Wesley.

TV Sports Markets (2006). European TV and Sports Rights 2006. London. TV Sports Markets.
Utt, J. & Fort, R. (2002). Pitfalls to Measuring Competitive Balance With Gini Coefficients. *Journal of Sports Economics, 3,* 367-373.
Vrooman, J. (1995). A General Theory of Professional Sports Leagues. *Southern Economic Journal, 61*, 971-990.
Vrooman, J. (1996). The Baseball Players' Labor Market Reconsidered. *Southern Economic Journal, 63*, 339-360.
Vrooman, J. (1997). A Unified Theory of Capital and Labor Markets in Major League Baseball. *Southern Economic Journal, 63*, 594-619.
Vrooman, J. (2000). The Economics of American Sports Leagues. *Scottish Journal of Political Economy, 47*, 594-619.
Walker, B. (1986). The Demand for Professional League Football and the Success of Football League Teams: Some City Size Effects. *Urban Studies, 23*, 209-219.
Walvin, J. (2000). *The People's Game: The History of Football Revisited.* Second Edition. Edinburgh, Scotland and London, UK: Mainstream.
Whitney, J. (1988). Winning Games versus Winning Championships: The Economics of Fan Interest and Team Performance. *Economic Inquiry, XXVI*, 703-724.
Wilson, P. & Sim, B. (1995). The Demand for Semi-Pro League Football in Malaysia 1989-91: A Panel Data Approach. *Applied Economics, 27*, 131-138.
Zimbalist, A. (1992a). *Baseball and Billions*. New York, NY: Basic Books.
Zimbalist, A. (1992b). Salaries and Performance: Beyond the Scully Model. In P. Sommers (Ed.), *Diamonds Are Forever: The Business of Baseball.* Washington DC: Brookings Institution Press.
Zimbalist, A. (2002). Competitive Balance in Sports Leagues: An Introduction. *Journal of Sports Economics, 3*, 111-121.
Zimbalist, A. (2003a). Competitive Balance Conundrums. Response to Fort and Maxcy's Comment. *Journal of Sports Economics, 4*, 161-163.
Zimbalist, A. (2003b). Sport as Business. *Oxford Review of Economic Policy, 19*, 503-511.