UNIVERSITY OF NOTTINGHAM

Discussion Papers in Economics

Discussion Paper
No. 00/6

THE ECONOMICS OF PROFESSIONAL SPORTS LEAGUES: A BARGAINING APPROACH

by Spiros Bougheas and Paul Downward

March 2000 DP 00/6
ISSN 1360-2438

UNIVERSITY OF NOTTINGHAM

Discussion Papers in Economics

Discussion Paper
No. 00/6

THE ECONOMICS OF PROFESSIONAL SPORTS LEAGUES: A BARGAINING APPROACH

by Spiros Bougheas and Paul Downward

Spiros Bougheas is Lecturer, School of Economics, University of Nottingham, and Paul Downward is Reader, Business School, Staffordshire University

March 2000

Abstract

This paper develops a theoretical framework of professional sports leagues which captures the bargaining process of sporting labour market negotiations. A

transfer market, which does not have a counterpart in other labour markets but it is an implicit and ubiquitous component of the market for professional players, finds a place within this framework. The effects of acquisitions of new players and transfers on the league's competitive balance and revenue distribution are examined and the implications of the results for league management policies, on both sides of the Atlantic, are analysed.
JEL J44, J50.

Acknowledgements: Helpful comments and suggestions from Richard Disney and Tim Worrall are greatly appreciated. We are responsible for any remaining error

---

*Corresponding Author: Economics Division, University of Nottingham, University Park, Nottingham, NG7 2RD; tel: 0044-(0)115-8466108; fax: 0044-(0)115-9514159; e-mail: spiros.bougheas@nottingham.ac.uk.

The Economics of Professional Sports Leagues: A Bargaining Approach

1. Introduction

Walter C. Neale's (1964) exposition of 'the peculiar economics of sports' helped to focus attention upon two important and closely related themes in the economic analysis of professional team sports. The first theme concerns an adequate economic conception of sporting leagues and the second concerns a clear statement of the 'uncertainty of outcome hypothesis'. More specifically, Neale made the first attempt to define a sporting league in economic terms and argued that leagues were examples of multi-plant monopolies being singly responsible for the administration of sports. For example, a team cannot determine its own output level (as measured by the number of games). That depends on how many teams the league admits and on how many times the league decrees teams play one another during the season. For Neale (1964), a central reason why leagues emerged as monopolies was the 'Louis-Schmelling Paradox'. Along with Rottenberg (1956), this was one of the first formal statements of the uncertainty of outcome hypothesis, which states that sports-fan interest is greatest when sporting competition is at its most intense. While each team strives to outperform its competitors, both on the field and financially, the league's survival depends on the maintenance of a competitive balance. In sports, every act of production requires the presence of an opponent – hence the paradox.

In contrast to Neale, Sloane (1971), in an analysis of European football, argued that a sporting league and its constituent teams may be more accurately viewed as a cartel than as a multi-plant firm. Teams in most sports take decisions about investment, about whether to produce at all, and they usually have substantial control over ticket prices. They also control merchandising which is now a major source of revenue. One of the most important aspects of Sloane's definition of sporting leagues is that it provides a more plausible framework within which to explore the potential problems associated with managing the mutual interdependence between teams. Neale's definition could imply that centralised control coordinates passive teams.

The problem of mutual interdependence can be illustrated in the context of league management policies. Many policies of cross-subsidisation have been justified on the basis of the uncertainty of outcome hypothesis. Leagues have attempted to transfer resources from stronger, more successful teams, to smaller less successful teams to try to initiate a 'virtuous circle' of overall benefit to the league that can be actively promoted and sustained by cross-subsidisation policies. The outcome of these policies has been weak. Thus, theoretical attempts to examine the effectiveness of cross-subsidisation policies in improving competitive balance have been developed. They have their origin in a cartel

model of sporting leagues developed by El Hodiri and Quirk (1971)[1]. These models take as their starting point a situation of competitive imbalance and examine the consequences of changes in league policies. Based on predictions developed from this model, Fort and Quirk (1995) provide some statistical inferences to support the view that these policies have been ineffective in the US. Quirk and Fort (1992) also provide detailed descriptive historical statistics to suggest that US Leagues tend to conform to long-run domination. In the UK Dobson and Goddard (1995) provide similar findings. Though the literature is somewhat scant this evidence can be seen as offering support for the 'invariance proposition'. This was originally suggested by Rottenberg (1956). It states that a league's competitive balance (talent distribution) is independent of the allocation of the ownership right to sell players services. It is a direct application of Coase's theorem.

A common characteristic of the modeling approach in this literature is the assumption that talent is homogeneous and perfectly divisible and, therefore, well suited for marginal analysis. An immediate consequence of this approach is that a team is identified, apart from some exogenous characteristics, by the total amount of talent it possesses. Furthermore, the price of talent (salaries) is determined in a competitive market[2].

In this paper, an alternative theoretical framework is proposed which captures the bargaining process of sporting labour market negotiations. A team in our model is defined as a collection of tasks. Players are characterised by their task-specific skills, and the level of their skills. An important feature of the model is that even if the level of a player's skills is team independent, the player's expected productivity differs across teams. The reason is that teams, in contrast to typical firms, hire more than one player for each task. While a team's performance depends on the talent it uses on the playing field, each player's contribution depends on the time he is expected to participate. As new players enter the league and older players retire the league's distribution of talent changes. This creates opportunities for the transfer of players to teams where

---

[1] For example, see El Hodiri and Quirk (1971), Sloane (19721), Scully (1974), Szymanski and Smith (1997) and Vrooman (1995).

[2] Bargaining is only considered in the empirical literature; see Carmichael and Thomas (1993), Dobson and Gerrard (1999), Dobson and Goddard (1998), Hylan, Lage and Treglia (1996), Kahn (1993), Krautman and Oppenheimer (1994), MacDonald and Reynolds (1994) and Marburger (1994).

they are expected to make greater contributions to a team's success. Thus, a transfer market, which does not have a counterpart in other labour markets but is an implied and ubiquitous component of the market for professional players[3], finds a place within our framework. Finally, salaries and transfer fees are determined through bargaining between the interested parties.

In the following section the variety of cross-subsidisation policies implemented by sporting leagues are reviewed. An important distinction between US and UK labour markets is noted. Section 3 outlines a 'benchmark' model. Three propositions are derived. These suggest that competitive balance in sporting leagues will depend on talent distribution but that the invariance proposition holds. Section 4 relaxes some of the assumptions of the model to explore more realistic scenarios. Consistent with the evidence discussed above it is shown that a lack of competitive balance is likely. In section 5 it is shown conditions corresponding to, for example, European sports labour market could challenge the invariance proposition.

2. Cross-Susidisation Policies

There have been a large number of ways in which leagues have intervened in the management of clubs finances in order to promote cross-subsidisation between clubs. While the particular administrative details may vary, however, they have had two major targets; the sporting labour market and revenue distribution[4]. Targeting sporting labour markets as a means of cross-subsidising clubs is based on the idea that players salaries and wages comprise a large proportion of sporting clubs costs in both the US and Europe[5]. Policies that affect players will thus have a large direct financial effect on clubs. Secondly, and directly concerned with the uncertainty of outcome hypothesis, it is the players who ultimately affect a club's success or failure in matches. Consequently, both the resources of clubs as well as their results can, in principle, be affected through policies aimed at the labour market. Leagues have attempted to influence club

[3] For example, in English football transfer surpluses have recently amounted to more than 30 per cent of gate revenues; Dobson and Goddard (1997).

[4] For a more extensive discussion of the institutional details the reader is referred to Downward and Dawson (forthcoming).

[5] See for example, Scully (1989) and Szymanski and Kuypers (1999).

financing and results through implementing three major types of labour market policy: Drafting systems, Salary caps, and Reserve option arrangements.

The most well known example of a drafting system is the 'Rookie Draft' in American football. Basically, a drafting system rations the order in which professional teams can sign new talent - rookies. In the reverse-order-of-finish draft of the NFL, teams that finish the lowest in the league get the first option to sign the best new talent.

While drafting systems, being targeted at the physical reallocation of sporting talent, have an indirect financial implication for clubs, salary caps are targeted at the financial cost of players directly. Salary caps imply a maximum amount that clubs can spend on players. One of the first examples in US sports was the NBA in 1980. The NFL adopted a cap in 1993 and baseball in the mid-1990s. In these cases clubs salary bills have been restricted to a certain proportion of clubs turnover. A similar policy in U.K. is currently in force in Rugby League and began in Rugby Union in 1999. In contrast, maximum wages for individual players were in force in football in the UK between 1900 and 1961. The intended implication and justification for these policies is that, in principle, it makes the best talent affordable to all teams.

The final form of labour market policy employed by sporting leagues has been Reserve Option clauses. The most famous example of this is in baseball where a form of this contract has been in force since 1880 and currently still exists in a much weaker form for rookie players. These clauses essentially tied players to clubs for their lifetime by giving the club the option to renew the player's contract when it expired. The retain-and-transfer system in European soccer is another example.

The other main form of cross-subsidisation policy adopted in sporting leagues has been to enforce clubs to redistribute some of their revenues. As far as gate-sharing arrangements are concerned, from its inception Baseball operated a 50-50 split on gate revenues. Half of gate receipts went to the home team and the remainder to the away team. However, away team shares have fallen steadily over the years and, currently, a variety of arrangements exist ensuring that the home team receives the largest share of revenues. Likewise, the NFL operates under a 60-40 split in favour of the home club while the NBA and NHL have no gate sharing. In Europe, similar arrangements have applied. In football for example, in the UK, between the 1920's and the 1980's an 80-20 split on gate revenues existed in favour of the home club. Moreover, the football league imposed a 4% levy on all receipts that were then redistributed in equal absolute shares. Similar arrangements existed, but were later abandoned, in rugby league.

As far as TV revenues are concerned, currently in the US, local TV coverage provides no revenue for visiting teams[6]. In contrast there are egalitarian arrangements for redistributing national TV revenues. The most radical changes in TV financing of sports has evolved outside the U.S. with the growth of BskyB and satellite TV coverage of sports in Europe and Australia. For example TV money has produced a huge financial gulf between those clubs within and those outside of the Premier League in Football or the Superleague in Rugby League (Baimbridge et al 1995, 1996; Symanski and Kuypers 1999). Moreover, the funding arrangements within these leagues further reinforce the financial gap between the successful and unsuccessful clubs.

These developments are an unusual phenomenon in economic terms. In any other industry the direct regulation of the terms under which firms compete in markets has been ruled as acting against the public interest. Despite the various legal frameworks for competition policy, in general, government intervention in sporting markets has been minimal. Indeed, in the case of Baseball in 1922, the Supreme Court in the US ruled that Baseball was exempted from the Sherman Act because it did not represent interstate commerce. While this decision has been criticised it has never been overturned. Moreover,

> "While other U.S. sports have not enjoyed the full exemptions granted to baseball, the competition authorities have still tended to look favourably on restrictive agreements. In particular the 1961 Sports Broadcasting Act exempted the collective selling of T.V. rights for sports leagues" (Symanski and Kuypers, 1999, p249).

However, one notable exception to this attitude has been in sporting labour markets with the rise of Free Agency. The move towards free agency in US sports began in the 1970's and by the 1990's the situation has developed to the extent that, for example, in baseball 'rookies' of 1 or 2 years experience are subject to a reserve clause. Intermediate players of between 3-6 years experience are eligible for final offer arbitration and veterans of 7 or more years service are free agents or eligible for final offer arbitration.

It is interesting to note that there are some strong parallels between the market for professional players in European association football and the US sporting labour markets. In particular labour market restrictions have applied to both player remuneration and player mobility[7]. The abolition of the maximum wage

---

[6] Fort and Quirk (1995).

[7] Symanski and Kuypers (1999).

for example took place in the 1960's. In addition a number of amendments to the retain-and-transfer system have been made. The retain-and–transfer system controlled player mobility in association football. Only a player registered with the Football Association can play professional football. Because the registration is held by a club, historically, it could control the player' movements much in the same way as the reserve-option clause in US sports. At the end of a season, for example, a club could retain players if it wished, or let them leave. In principle it could retain the registration of a player even if it did not renew the contract. Moreover, clubs could charge a fee - a transfer fee - for allowing the player to move to another club. Note, that this could apply even in the absence of the maximum wage, so, effectively, the terms and conditions of the players' contracts lay with the club under this system.

The most celebrated amendment to the retain-and-transfer system has been the recent Bosman Ruling. The European Court of Justice, under Article 177 of the Treaty of Rome which enshrines the free mobility of land, labour and capital in the European Union, declared that, in the absence of pressing reasons of public interest, the transfer rules did constitute an obstacle to the free movement of workers. Thus, the important outcome of the Bosman Ruling was that *no* fee could be expected by clubs on the transfer of an out-of-contract player. It is interesting to note that, the Advocate General of the European Court of Appeal accepted, in principle, the need for Leagues to maintain competitive balance and uncertainty of outcome in making his ruling. He also accepted that smaller clubs often covered financial losses through transfer fee income. However, he argued that using the 'means' of the transfer system to achieve these 'ends' was not justifiable. This was because there were other methods of achieving competitive balance that did not restrict player mobility.

Of more importance to this paper is the fact that despite the general similarity of developments in European and US sporting labour markets, it remains that different specific contractual structures exist. In the U.K. contract duration remains potentially more flexibly determined compared to the US because of the latter threefold partition of the labour market. Indeed evidence in the US by Kahn (1993) supports the view that contract duration only increases after US athletes become eligible for free agency after 7 years experience. In the UK no such constraints exist. The potential effects of a more general increase in contract length are examined in section 5.

3. A Bargaining Model of Sporting Leagues: The Benchmark Case

In order to capture the current bargaining nature of sporting labour market negotiations and to get some insight into the dynamics of the transactions, especially transfers, that characterise sporting leagues, the following model is proposed. Consider a league consisting of N teams. The economic success of each team depends on both their relative performance against the rest of their league and the success of the league as a unit. Relative performance, which itself depends on some measure of relative talent, determines the team's share of league revenues. The level of league revenues, which provides a measure for the success of the league as a unit, depends not only on the aggregate level of talent but also on its distribution. For example, if a single team dominates the league for a long period of time then it will reduce both competitive balance and outcome uncertainty, thus, reducing the league's aggregate revenues.

Each team allocates talent to M tasks (playing positions) by hiring players. We assume that each player's talent is suitable for only one task and it takes one of three possible values. There is an abundant supply of players with only basic skills, $\underline{t}$. The rest of the players are either low, t_l, or high, t_h,, talented; where: $t_l<t_h$. Since the players who are either low or high talented are in short supply only the basic skill level is guaranteed for each position. Moreover, even if a team hires players who belong to the top two talent groups, it is assumed that these players will be on the playing field with probability p. Because of this restriction, teams might find it profitable to hire more than one talented player for each task. For simplicity, we restrict the number of talented players allocated to each position to be less than or equal to two. Of course, the team's performance depends on the amount of talent used and not on the total amount of talent it possesses.

In the benchmark case, both team and league revenues are proportional to the talent in use. Therefore, if a team hires a talented player then both its own and the league's expected revenues are raised by the same amount. After an appropriate normalization of talent units we can set revenues identically equal to talent in use. Let t_i^j denote the talent of team i at task j. Then the total talent (revenues) of team i, t_i , is given by:

$$(1) \qquad t_i = \sum_{j=1}^{M} t_i^j$$

and the league's total talent (revenues) is given by:

$$(2) \qquad t = \sum_{i=1}^{N} t_i = \sum_{i=1}^{N} \sum_{j=1}^{M} t_i^j$$

Team costs are equal to the amount paid for players salaries. The evidence suggests that the market for basic skills – or rookies – is monopsonistic. It is assumed that wages are proportional to talent and after a normalisation are set equal to $\underline{t}$. In contrast, the salaries of talented players are determined by a bargaining process between them and their employers. Contracts are legally binding and can only be broken under mutual consent. Recently, Stole and Zwiebel (1996) developed a framework for analyzing intra-firm bargaining when contracts are non-binding. In their set-up firms and workers commit only to a wage conditional on employment. Before production begins any party can initiate a contract renegotiation. As new opportunities arrive, they alter the outside options of the two parties and can have significant effects on the employment relationship. In contrast, in the present model, because contracts are binding, such considerations are irrelevant. Assuming Nash bargaining, the two parties split the surplus generated by hiring the player.

Derivation of Surplus

When a team hires a second talented player the surplus generated by that player depends not only on his / her talent but also on the talent of the player hired before him / her. There are seven cases to be considered that depend not only on the type of players hired by the team but also on the order of hiring. The first column of table 1 describes these seven cases. For example, row 5 corresponds to the case where at task *j*, team *i* has one low talented player and a one high talented player and the low talented player was hired first. The second column shows the talent of team *i* at task *j* and the third column the surplus generated by the last player hired. Because in the benchmark model total team talent is equal to the sum of talents at each task the analysis will focus on a single task.

The derivation of the expressions in the second column is based on the following two considerations. First, what matters for team performance is the talent used in competition and not the available talent. Second, even if a player is needed, he / she will be on the playing field only with probability *p*. The first row refers to the case where the team does not have any talented player available at this particular position. Given that there is an unlimited supply of players with basic skills, the team's talent at this task is $\underline{t}$. The following two rows refer to the case where the team has only one talented player. This player is expected to be playing with probability *p* and, therefore, with probability *1-p* the team will have to replace him / her with a player with only basic skills. The next four rows refer to the case where the team has two talented players. If there is one low talented player and one high talented player, the team's first choice is the high talented player. In the case described in row 5 the team first hired a low talented player and then a high talented player. The high talented is expected to play a fraction *p* of total playing time and the low talented player a fraction *p* of the remaining

time. Of course, for the calculation of total talent the sequence of hiring is irrelevant, hence, rows 5 and 6 show the same expected talent in use.

For the calculation of expected surplus, the sequence of hiring is important. The expected surplus generated by a newly hired player depends on players hired before him/her. This player might also affect the expected surplus of players hired before him/her. However, because contracts are binding the team cannot renegotiate existing contracts. Therefore, the new player's salary will be based on the surplus he/she is expected to generate and not necessarily on his/her absolute talent. For example, to calculate the expected surplus in row 5 subtract the talent in use shown in row 2 from the one shown in row 5. The difference captures the additional expected talent in use which a high talented players offers to a team which already possesses a low talented player in the same position. On the contrary, in the framework of Stole and Zwiebel (1996), the team would renegotiate the low talented player's contract so that the new contract would reflect his/her diminished contribution caused by hiring the high talented player. Notice that the expected surpluses shown in rows 4 and 6 are equal even if the expected talents in use are not. This is because the team hires a second player who is low talented and, consequently, he/she will be on the playing field with probability $p(1-p)$[8].

Bargaining, Salaries and Profits

The salary of a talented player is determined by a Nash bargaining process between him/her and the team owner. Both parties understand that any agreement signed is final and can only be amended under mutual consent. The bargaining outcome is a split of the expected surplus. Let γ denote the fraction received by the player, which is assumed to be uniform across both teams and players. Without any loss of generality, normalise basic skills to zero. Then, the expected cost of team i, C_i, is given by:

$$C_i = \gamma t_i \quad (3)$$

and its expected profits, are given by $(1-\gamma)t_i$. Despite its simplicity, the above model is sufficiently rich and captures these characteristics of sports leagues that set them apart from other sectors of the economy. Next, we consider a simple dynamic version of the benchmark model and examine the effects of transactions in sporting labour markets.

---

[8] It is assumed that if the newly hired player has the same talent as the player hired before him/her then he/she will be asked to play only if the other player is not available.

Dynamic Considerations

Most of the important issues related to professional team sports are dynamic in nature. The conflict between intra-league competition and the survival of the league itself is, essentially, a conflict between the short-run interests of individual teams and the long-run concern of all teams to survive as a unit. In this model time is divided in discrete periods. Each period represents a full season and is divided into two sub-periods. The first sub-period corresponds to preseason where teams make their personnel decisions. During the second sub-period teams compete during the 'season'. For simplicity, acquisition of talented players is not allowed during the season.[9]

At the beginning of each period, each team inherits the talent they finished with, one period earlier[10]. This defines a beginning of period distribution of talent for each team at each task. To complete the dynamics the effects of the preseason transactions on the above distribution need to be specified. It is assumed that a small number of talented players enter the league during each preseason. They are free to negotiate a contract with any team and there is sufficient time to contact all teams. If two teams make the same contract offer then the employment decision is decided randomly. For simplicity, it is assumed that new talented players enter the market sequentially and only after the new players before them have completed their negotiations. It is further assumed that the inflow of players is quite uncertain[11]. Of course, with no additional restrictions, no matter how slow the process is, eventually, all teams will be able to fill all positions with two high talented players. Therefore, it is further assumed that a small number of players leave the league at the end of each season. Both the talents and the positions of incoming and outgoing players is random. At this point, the reader might feel uncomfortable about the lack of a rigorous specification of the model's dynamics. However, as it will be demonstrated below, none of the results of the benchmark model depend on any

---

[9] Basic skill players might be needed for replacement.

[10] Below this assumption is relaxed so that the effect of free agency can be examined.

[11] The reason behind this assumption is to simplify the owner's decision process. More specifically, when the opportunity to sign a new low talented player comes they will have the incentive to do so rather than wait for a better one.

particular specification of the initial talent distribution or on the distributions that specify the entry of the new players and exit of those leaving the league.

New Talent and Competitive Balance

As new players enter the league and sign contracts they alter the talent distribution. If, as a result of these new contracts, the distribution gets less uniform over time then the league's competitive balance will be destroyed. In contrast, for the benchmark case the following proposition holds[12]:

Proposition 1: In the benchmark model, entry of new players, on average, increases the league's competitive balance.

The intuition is straightforward. The surplus that new talented players expect to generate, and hence the salaries they expect to receive, is higher the lower the talent their new team possess at their task. Therefore, on average, low talented teams, have a better chance of signing new talented players.

The Transfer Market and Competitive Balance

The transfer market is a distinct characteristic of sporting labour markets. In contrast to firms in other labour markets, sports teams hire more than one player for employment in the same task. It has been assumed that talent is exogenously given. However, because contracts are legally binding, each player's expected surplus depends on the talent of the player hired before him/her and consequently, a player's expected surplus can be different across teams. Then, a transfer market can be beneficial to both teams and players if the surplus the players are expected to generate with their new teams is higher than the one they produce at their old teams. The surplus difference will be referred as the net expected surplus of the transfer. The player receives a fraction of the surplus and the rest is divided between the two teams. The payment received by the selling team is called the transfer fee. Teams might also find it beneficial to exchange players if such an exchange generates a positive net expected surplus. The following proposition, refers to the effects of the above transactions on the league's competitive balance.

Proposition 2: In the benchmark model, transfers and exchanges of players, on average, increase the league's competitive balance.

[12] Propositions 1 and 2 are proved in the Appendix.

Again, the intuition is simple. These type of transactions involve talented players who move from teams with other talented players at their specialized task, and hence their expected surplus is low, to teams with lower talent at the same task. On average, it is expected that low talented teams have a better chance of benefiting from these transactions.

Propositions 1 and 2 are quite robust since their proof does not require any restrictions on the distribution functions or the values of the model's parameters.

Free Agency and Competitive Balance

Does the invariance proposition hold in the benchmark case? In order to address this question the assumption that at the beginning of each period teams inherit the talent they have finished with one period earlier must be relaxed. Suppose that at the end of each season players are free to move to another team. Such a position is consistent with free agency[13]. It is clear, that a player will decide to move to another team only if his expected surplus at his/her new team is higher. Therefore, transfers do not depend on who owns the right to sell the player's services. The only difference is that under free agency it is only the player and his/her new team's owners who share the surplus; i.e. there is no transfer fee. The following proposition summarizes this argument:

Proposition 3: In the benchmark model, the league's competitive balance is independent of the ownership of the right to sell players services.

4. Explaining competitive imbalance

The benchmark model has provided a simple analytical framework for addressing some of the issues concerning competitive team sports leagues. In comparison with the rest of the literature, the modeling approach followed in this paper has offered a variety of novel features. It has completely dispensed with the assumptions that talent is homogeneous, divisible and inelastically supplied and, therefore, has allowed the introduction of a bargaining approach to salary determination which is in accord with experience. In addition, observation of common practice has dictated the postulate that contracts are

[13] It is implicitly assumed that players sign one-period contracts. The issue of optimal contract length is discussed below.

legally binding which, as it was argued above, in conjunction with the fact that teams hire more than one player at each task, offers a rationale for the emergence and extensive use of the transfer market. It has been demonstrated that the benchmark model is sufficiently rich to yield predictions about the effects of transactions in the market for talent on the league's competitive balance and, furthermore, we have managed to derive a simple version of the invariance proposition.

Despite its merits, the benchmark model has a linear structure that is definitely restrictive. For both analytical and expositional simplicity, the following assumptions were made: a) a team's expected talent in use is equal to the sum of the expected talents in use at each task, b) each team's share of league revenue is proportional to its share of league talent, and c) the league's total revenue is proportional to the league's total talent. In this section of the paper we relax these assumptions and discuss teams' access to capital markets. We find that there are strong reasons why competitive imbalance might be expected in sporting leagues.

Access to Capital Markets

Thus far, we have ignored the presence of costs other than salary payments. This is the case, because implicitly it has been assumed that teams have access to perfect capital markets. In this case, even if a team's other costs set limits on its liquid assets, it could still finance the acquisition of new players either through bank lending or by issuing more shares. However, in reality, teams, especially those located at the bottom of the league tables, have limited access to capital markets. The reason is the uncertainty surrounding the talent of new players which is partially captured by the parameter p in the benchmark model. It is hard to believe that team owners and potential lenders have the same beliefs about the talents of new players and such informational asymmetries can limit the ability of teams to raise outside funds. If that is the case then the presence of other costs can inhibit the access of poor teams either to the market for new players or to the transfer market. If players move only to teams that can afford them then the league's competitive balance could be destroyed.

Relaxing the model's assumptions

A team's expected talent in use is equal to the sum of the expected talents in use at each task. In contrast to the rest of the literature, the model has allowed for a multiplicity of tasks in order to capture the idea that players are not perfect substitutes for each other. However, the first assumption that a team's expected

talent in use is equal to the sum of its corresponding talents at each task implies that tasks are perfect substitutes. While perfect substitutability is not crucial for the conclusion of Propositions 1 and 2 that, on average, transactions in the sporting labour market improve the league's competitive balance, there is no doubt that the higher the degree of substitutability the more robust the conclusion is. For example, consider the opposite extreme where tasks are perfect complements which implies that team's *i* talent is given by:

$$(4) \qquad t_i = \min\{t_i^1,...,t_i^j,...,t_i^M\}$$

This corresponds to the case where teams need to be well balanced across all tasks in order to be competitive. Suppose that one team has filled all positions with high talented players with only one exception, task *j,* where it has only a low talented player. Another team has filled all positions with low talented players except task *j* where there is a high talented player available. Then an exchange of players can improve the first team without affecting the absolute strength of the second team. Competitive balance in this case is harmed. While the two teams have started with the same talent in use, after the transaction one of the teams has become considerably stronger.

Each team's share of league revenue is proportional to its share of league talent. This is a crucial supposition. Consider an equally plausible case where successful teams tend to receive a disproportionately high share of league revenues. Then, transactions in the market for new players and in the transfer market might destroy the league's competitive balance. At this point, it is important to distinguish between the surplus in talent that a new player is expected to generate from the revenues that this surplus is expected to yield. A high talented player is expected to generate a higher surplus of talent with a team that has, for example, only players with basic skills at his/her specialized task than with a team that has another talented player. However, his/her talent might be more profitable for the second team if the latter is, overall, more successful. Similarly, a relatively weak team might sell its high talented players to stronger teams if such transfers generate positive expected revenues[14]. As Downward and Dawson (forthcoming) note, to the extent that demand, and hence individual team revenues, depend on a team's success rather than uncertainty of outcome in the league, this might provide the conditions for such a scenario. Further, the limited access to capital markets discussed earlier might promote the sale of better players by poorer teams.

[14] The case of task complementarity, discussed above, is a good example.

The above discussion suggests the following important observation. Transactions can affect the rest of the league, something that is impossible within the benchmark case where transactions only affect participating teams. These externalities imposed on other teams might be strong enough that the overall effect of transactions in the sporting labour market could be deterioration in the league's competitive balance – in contrast, to the conclusions drawn from Propositions 1 and 2. On the contrary, the invariance proposition (Proposition 3) still holds in the more general case. Once more, the allocation of the right to sell the players services affects only the distribution of the revenues they are expected to yield and not the distribution of talent in the league.

The league's total revenue is proportional to the league's total talent. While the first two suppositions of the benchmark model are the most analytically convenient among many other equally plausible scenarios, the final supposition that total league revenues are proportional to total team talent is too strong. In fact, it can be argued that according to the uncertainty of outcome hypothesis that total league revenues should be higher the more uniform the distribution of talent across teams is. In other words, we need to consider moments of the distribution higher than the first. Nevertheless, doing so will only strengthen the observations made above with respect to the first two suppositions. Indeed, more generally, a monotonic relationship between competitive balance and league revenues will suffice. Consequently, the results above do not necessarily hinge on the uncertainty of outcome hypothesis.

5. Youth Team Policies: A Challenge to the Invariance Proposition

As discussed in section 3, while there have been some similar developments in European and US sporting labour markets, a key difference is the potential for European sporting contract length to vary more than in the US. An expressed implication associated with the ruling was that contract duration would increase.[15] Indeed, since the Bosman ruling, a number of European football clubs have signed consistently long-term agreements with their young players[16]. In the context of the previous analysis, while the above observation can be explained by the ability of team owners to extract a share of the surplus generated by players moving to other teams, the invariance proposition makes clear that such transfers do not affect the league's distribution of talent. In this final section,

[15] Sir John Hall, Chairman of Newcastle United Football Club PLC, quoted in the Times 20/9/95.

[16] For example, Manchester United.

another plausible explanation that challenges the very premises of the invariance proposition is considered.

The Coase theorem that lies at the heart of the invariance proposition holds in environments where transaction costs are absent and, hence, parties have the ability to write complete contracts. In recent years, the fast developing literature on incomplete contracts[17] has been addressing problems arising when parties are unable to write contracts contingent on all future events that might affect their relationship. This inability might stem, for example, from the fact that certain actions taken by the two parties cannot be verified by a third party (e.g., court of law). If as a consequence of these actions there are rents to continuing the relationship then, unless the two parties can make long-term commitments, these actions might be taken at sub-optimal levels.

These ideas have been extensively explored in the labour economics literature[18]. Firms often make investments that enhance the future productivity of their workers. In many cases it is not feasible to specify all the characteristics (e.g. quality) of these investments in sufficient detail so that can be verified by a third party. In the absence of long-term agreements this might lead to underinvestment, especially if these skills are useful to other firms[19]. If workers are not financially constrained then it is optimal for the workers to finance their training. However, because investments are not verifiable it is doubtful that workers can raise funds in the financial markets.

The market for professional sports team players offers a straightforward application of these ideas. The aim of youth team policies is to improve the skills of young players. The implementation of these policies requires costly investments in training grounds, other related facilities and professional staffing. Teams expect that the benefits of their investment will be realized when the players mature and become regular first team members. However, at that time other teams might attempt to attract these players and, thus, capitalize on another team's investment. At this point, the issue of verifiability becomes crucial. It might be impossible for courts to verify that a player's talent is the result of the team's youth policy. As a consequence, team owners might be reluctant to invest in these policies if they believe that they might not be compensated.

---

[17] See Hart (1995).

[18] See Malcomson (1997) for a comprehensive review of this literature.

[19] See Stevens (1996) and Adnett and Bougheas (1998) for the case of transferable training.

A long-term binding contract, effectively, transfers to teams the ownership of the right to sell the players services. Teams and players can agree on a salary profile which overtime compensates teams, thus, providing them with the right incentives to invest. Furthermore, in the case of a transfer, the team can capture part of the surplus, since the original contract can only be broken by mutual consent. In contrast to the invariance proposition (Proposition 3), the foregoing discussion clearly suggests that the league's talent distribution depends on the allocation of control rights because the latter have a direct effect on teams' decisions to invest in youth training policies. Significantly, to the extent that US sporting labour markets differ from their European counterparts this suggests the potential for a test of the invariance proposition. More generally, the discussion reinforces the view that the specific ways in which sporting leagues evolves may well matter in understanding the development of competitive balance. Consequently league-specific research that captures the institutional realities of particular sporting leagues may yield interesting insights into this development.

6. Conclusion

This paper has suggested a novel theoretical framework of sporting leagues that more adequately captures the bargaining process characterising sporting labour markets. Along with the rest of the literature, it has been argued that if an unrestricted transfer market exists and capital markets are perfect then free agency does not threaten competitive balance, which is a restatement of the invariance proposition. However, it was also stressed that the invariance proposition does not imply that transactions in the market for professional players cannot have destabilising effects. The multi-team league model clearly suggests that from the moment there is a digression from the benchmark case then the above transactions can have significant effects on the distribution of talent across teams. Abandoning the linear structure of the model implies that transfers, acquisitions and exchanges of players can impose negative externalities on teams not participating in these transactions with potential adverse consequences for the league's competitive balance. Furthermore, imperfect capital markets can limit the ability of poor teams to finance the acquisition of players. However, attempts to mitigate these effects by policies that redistribute revenues (cross-subsidisation) can distort the incentives of firms 'near the bottom' to compete. Finally, it has been argued that in order to understand the recent proliferation of long-term contracts in European football it is important to appreciate that the invariance proposition ceases to hold in an incomplete contacting environment.

References

Adnett, N. and Bougheas, S. (1998) “Transferable Training and the Hold-Up Problem: Policy Implications,” Paper Presented at the Meetings of the European Economic Association (Berlin, September 1998).

Baimbridge, M., Cameron, S. and Dawson P. (1995) 'Satellite Broadcasting and Match Attendance: The case of Rugby League', *Applied Economics Letters*, 2, 10: 343--46.

Baimbridge, M., Cameron, S. and Dawson, P. (1996) 'Satellite Television and the demand for football: A whole new ball game?' *Scottish Journal of Political Economy*, 43, 3: 317—33

Carmichael, F. and Thomas, D. (1993) "Bargaining in the Transfer Market: Theory and Evidence," *Applied Economics* 25, 1467-76.

Dobson, S. and Gerrard, B. (1999) 'The Determination of Player Transfer fee's in English Professional Soccer' *Journal of Sport Management*, Vol 13, No 4, pp259-274.

Dobson, S.M. and Goddard, J.A. (1995) 'The demand for professional league football in England and Wales, 1925-92', *Discussion Paper, Department of Economics,* University of Hull.

Dobson and Goddard (1998) 'Performance, Revenue and Cross Subsidisation in the Football League 1927-1994' *Economic History Review*, LI, 4, pp763-785.

Downward, P. and Dawson, A. The Economics of Professional Team Sports (forthcoming) Routledge.

El Hodiri, M. and Quirk, J. (1971) "An Economic Model of a Professional Sports League," *Journal of Political Economy* 79, 1302-19.

Fort, R. and Quirk, J. (1995) "Cross-subsidization, Incentives, and Outcomes in Professional Team Sports Leagues," *Journal of Economic Literature* 33, 1265-99.

Hart, O. (1995) Firms, Contracts and Financial Structure, Oxford: Clarendon Press.

Malcomson, J. (1997) "Contracts, Hold-Up, and Labour Markets," *Journal of Economic Literature* 35, 1916-57.

Hylan, T., Lage, M. and Treglia, M. (1996) "The Coase Theorem, Free Agency and Major League Baseball," *Southern Economic Journal* 62.

Kahn, L. (1993) "Free Agency, Long-term Contracts and Compensation in Major League Baseball: Estimates from Panel Data," *The Review of Economics and* Statistics 75, 157-64.

Krautmann, A. and Oppenheimer, M. (1994) "Free Agency and the Allocation of Labor in Major League Baseball," *Managerial and Decision Economics* 15, 459-69.

MacDonald, D. and Reynolds, M. (1994) "Are Baseball Players Paid Their Marginal Products?" *Managerial and Decision Economics* 15, 443-57.
Marburger, D. (1994) "Bargaining Power and the Structure of Salaries in Major League Baseball," *Managerial and Decision Economics* 15, 433-441.

Neale, W. (1964) "The Peculiar Economics of Professional Sports," *Quarterly Journal of Economics* 78, 1-14.

Quirk J.P. and Fort, R., Pay Dirt: The Business of Professional Team Sports, Princeton N.J.: Princeton University Press, 1992.

Rottenberg, S. (1956) "The Baseball Players' Market," *Journal of Political Economy* 64, 242-58.

Scully, G. (1974) "Pay and Performance in Major League Baseball," *American Economic Review* 64, 915-30.

Scully, G., The Business of major League Baseball, Chicago:University of Chicago Press, 1989.

Sloane, P. (1971) "The Economics of Professional Football: The Football Club As A Utility Maximizer," *The Scottish Journal of Political Economy* 18, 121-46.

Stevens, M. (1996) "Transferable Training and Poaching Externalities," in Booth, A. and Snower, D. (eds.) Acquiring Skills: Market Failures, Their Syptoms and Policy Responses, Cambridge: Cambridge University Press.

Stole, L. and Zwiebel, J. (1996b) "Intra-firm Bargaining under Non-binding Contracts," *Review of Economic Studies* 63, 375-410.

Syzmanski, S. and Kuypers, T., Winners and Losers: The Business Strategy of Football, Viking Press:Harmondsworth, 1999.

Szymanski, S. and Smith, R. (1997) "The English Football Industry: Profit, Performance and Industrial Structure," *International Review of Applied Economics* 11.

Vrooman, J. (1995) "A General Theory of Professional Sports Leagues," *Southern Economic Journal* 61, 971-90.

Appendix

Proof of Proposition 1

It is demonstrated that, on average, new players sign contracts with teams which are relative weaker at their task of specialization. There are three cases we need to consider for each type of talent. A new player will either sign a contract with a team which does not have any other talented player in the same position or with

a team which has one more talented player who, in turn, is either high or low talented.

For new low talented players the relevant cases are shown in rows 2, 4, and 6 of table 1. The expected surplus in cases 4 and 6 are equal and lower than the expected surplus in case 2. A low talented player prefers to sign a contract with a team that has only players with basic skills available at the same position. In such a team he/she is expected to play a fraction *p* of the time, in contrast, in the other two cases he/she is expected only to play as a substitute, i.e. a fraction *p(1-p)* of the time. If the player's only choice are teams which have another talented player in the same position then the player is indifferent between the two alternatives and the decision is determined randomly.

For high talented players the corresponding cases are shown in rows 3, 5, and 7 of table 1. Subtracting the expected surplus shown in row 5 from the one shown in row 3, we find that the difference is equal to $p^2(t_l - \underline{t}) > 0$. Therefore, a new high talented player prefers to sign a contract with a team that has only players with basic skills available at the same position rather than with a team that has another low talented player. Similarly, we find that a high talented player prefers to sign a contract with a team that has another low talented player rather than with one that has another high talented player. Subtracting the expected surplus shown in row 7 from the one in row 5, we find that the difference is equal to $p(1-p)(t_h - t_l) > 0$. A new high talented player is expected to play a fraction *p* of the time when the other player is low talented and only as a substitute if the other player is also high talented.

Up to this point, it has been shown that new players reduce team differences in talent at their playing position. It is still possible that a new player might decrease the league's competitive balance by signing with a team that has low talent in his/her position but high talent in other positions. However, the skill of each new player is randomly determined and, therefore, teams which have talent deficiency in many positions, i.e. low average talent, have a greater chance of matching these positions with the skills of new players. It is the last consideration that warrants the qualifier 'on average' in the statement of the Proposition.

Proof of Proposition 2

It must be shown that all transfers and exchanges of players that generate a positive net expected surplus, on average, increase the league's competitive balance.

Consider transfers. Notice, that symmetric transfers, i.e. transfers that leave the league's talent distribution intact, generate zero expected surplus and, therefore, can be ignored. For example, a transfer of a talented player from a team that has not another talented player in the same task to a team with no talented players (in the same task) generates zero expected surplus.

Any transfer of a talented player from a team with another talented player in the same task to a team with no talented players (in the same task) generates positive expected surplus and increases the league's competitive balance. This follows from Proposition 1. It has been previously established that the expected surplus of talented players in teams with another talented player in the same task is lower than their expected surplus in teams with no talented players (in the same task).

Finally, consider transfers of talented players from teams with two talented players to teams with only one talented player in the same task. The only such transfer generating a positive expected surplus is of a high talented player from a team with two high talented players to a team with only a low talented player in the same task. The only such transfer that generates a positive expected surplus is when a high talented player moves from a team with another high talented player to a team with one low talented player. To find the expected surplus of this transfer subtract the expected surplus shown in row 7 of table 1 from the one shown in row 5. The difference is equal to $p^2(t_h - t_l) > 0$. It must be also shown that the transfer improves competitive balance. Before the transfer the difference in expected talent in use was equal to $p(t_h - t_l) + (1-p)t_h - p(1-p)\underline{t}$ (subtract the expected talent in use shown in row 2 from the one shown in row 7). After the transfer the difference is equal to $p(1-p)t_l - p(1-p)\underline{t}$ (subtract the expected talent in use shown in row 3 from the one shown in row 5). Since $t_h > t_l$ the difference has been reduced.

Next, consider exchanges. There is one case that generates a positive expected surplus. A team with two high talented players exchanges one of them for one of the two low talented players of another team. The exchange reduces the expected talent in use of the team which originally had two high talented players by $p(1-p)(t_h - t_l)$ (subtract the expected talent in use shown in row 6 from the one shown in row 7) while it increases the expected talent in use of the team which originally had two low talented players by $p(t_h - t_l)$. It is clear that the overall expected change is positive. Furthermore, after the exchange the two teams have
equal talent in use in that task.

For the same reasons given in the proof of Proposition 1 the qualifier 'on average' has been added in the statement of Proposition 2.

Table 1: Talent in Use and Expected Surplus in the Benchmark Model

Available Players	*Expected Talent in Use*	*Expected Surplus*
$\{\underline{t}\}$	$\underline{t}$	
$\{t_l\}$	$pt_l + (1-p)\underline{t}$	$p(t_l - \underline{t})$
$\{t_h\}$	$pt_h + (1-p)\underline{t}$	$p(t_h - \underline{t})$
$\{t_l, t_l\}$	$p(2-p)t_l + (1-p)^2\underline{t}$	$p(1-p)(t_l - \underline{t})$
$\{t_l, t_h\}$	$pt_h + p(1-p)t_l + (1-p)^2\underline{t}$	$p(t_h - pt_l - (1-p)\underline{t})$

$\{t_h, t_l\}$	$pt_h + p(1-p)t_l + (1-p)^2\underline{t}$	$p(1-p)(t_l - \underline{t})$
$\{t_h, t_h\}$	$p(2-p)t_h + (1-p)^2\underline{t}$	$p(1-p)(t_h - \underline{t})$