The B.E. Journal of Economic Analysis & Policy

Topics

Volume 9, *Issue* 1 2009 *Article* 17

The Effect of Salary Caps in Professional Team Sports on Social Welfare

Helmut M. Dietl* Markus Lang†
Alexander Rathke‡

*University of Zurich, helmut.dietl@isu.uzh.ch
†University of Zurich, markus.lang@business.uzh.ch
‡University of Zurich, rathke@iew.uzh.ch

Recommended Citation
Helmut M. Dietl, Markus Lang, and Alexander Rathke (2009) "The Effect of Salary Caps in Professional Team Sports on Social Welfare," *The B.E. Journal of Economic Analysis & Policy*: Vol. 9: Iss. 1 (Topics), Article 17.

The Effect of Salary Caps in Professional Team Sports on Social Welfare*

Helmut M. Dietl, Markus Lang, and Alexander Rathke

Abstract

This paper provides a theoretical model of a team sports league and studies the welfare effect of salary caps. It shows that salary caps will increase competitive balance and decrease overall salary payments within the league. The resulting effect on social welfare is counter-intuitive and depends on the preference of fans for aggregate talent and for competitive balance. A salary cap that binds only for large-market clubs will increase social welfare if fans prefer aggregate talent despite the fact that the salary cap will result in lower aggregate talent. If fans prefer competitive balance, on the other hand, any binding salary cap will reduce social welfare.

KEYWORDS: salary caps, social welfare, competitive balance, team sports league

*We would like to thank Martin Bernhard, Egon Franck, Leo Kahane, Stefan Késenne, Wolfgang Köhler, Marco Runkel, Bernd Schauenberg, Oliver Williamson, Ulrich Woitek, conference participants at Magglingen 2007, and seminar participants from the Universities of Antwerp, Freiburg, Munich and Zurich for helpful comments and suggestions. We also gratefully acknowledge the financial support provided by the Swiss National Science Foundation (SNSF research project no. 105270) and the research fund of the University of Zurich.

1 Introduction

A salary cap is a limit on the amount of money a club can spend on player salaries. The cap is usually defined as a percentage of average annual revenues and limits the club's investment in playing talent. Since most leagues compute their caps on the basis of the revenues of the preceding season, the cap is actually a fixed sum. In 2006, for example, the US National Football League (NFL) had a salary cap of approximately 102 million US dollars per team.

The North-American National Basketball Association (NBA) was the first league to introduce a salary cap for the 1984-1985 season.[1] Today, salary caps are in effect in professional team sports all around the world. In North America, the National Hockey League,[2] the Canadian Football League, the National Football League, the National Basketball Association and the Arena Football League have all installed salary caps. In Australia, the Australian Football League, the National Rugby League and A-League Soccer have implemented salary caps to regulate their labor markets. In Europe, salary caps are in effect in the Guiness Premiership in rugby union and the Super League in rugby league. In European soccer, a small fraction of European football clubs, known as G 14 and established as an interest group of 18 prominent clubs of European football, had already brought up the issue of salary cost controls in 2004. The members of G 14 had planned to limit their salary expenditures at 70% of audited club turnover from the season 2005/2006 onwards. At the same time the minimum allowable amount for total staff costs of each member was set at 30 million Euros. According to the G 14 plan, verification of the clubs' compliance with these principles should be carried out by their statutory auditors. However, the G 14 plan has never been put into practice and G 14 dissolved in January 2008, when the new European Club Association was founded under the auspices of UEFA.[3]

From an economic perspective, salary caps are often regarded as a collusive agreement of wealthy owners to use their monopoly power to transfer player rents back to ownership.[4] Nevertheless, salary caps are not illegal in the US because they are the result of a freely-negotiated collective bargaining agreement between the players' union and the league, represented by their governing body. The stated rationale for salary caps focuses on two main

[1]See Staudohar (1998).

[2]A lockout in 2004-2005 resulted, for the first time, in the loss of an entire season in the National Hockey League. The main point of contention was that the club owners insisted on the introduction of a salary cap to have cost-certainty (Staudohar, 2005).

[3]See Késenne (2003) for an analysis.

[4]See e.g. Vrooman (1995, 2000).

objectives: increasing competitive balance and maintaining financial stability. The concern for competitive balance describes one of the most important peculiarities of professional team sports:[5] It is a wideheld belief that a certain degree of uncertainty about the outcome is necessary to ensure an entertaining competition.[6] Salary caps prevent large-market clubs from becoming too dominant by helping small-market clubs to keep star players who would otherwise be attracted by higher salary offers from large-market clubs. Fort and Quirk (1995) consider an enforceable salary cap as the only effective device to maintain "financial viability" and improve competitive balance.

In Europe, the leading football clubs cited the protection of the financial future of the game as the main reason for their attempts to introduce a salary cap. Many clubs are facing financial ruin after gambling on spiralling wages. Owing to its structure, professional team sport carries the risks that its clubs over-invest in playing talent (see Dietl et al., 2008). Salary caps prevent clubs from over-investing in playing talent.

Both arguments have been discussed in the economic literature. According to Rottenberg (1956) clubs would not voluntarily bid themselves into bankruptcy and diminishing returns in terms of talent will guarantee at least some level of competitive talent. Whitney (1993), on the other hand, shows that the market for star athletes in professional team sports is subject to destructive competition - a process which drives some clubs into bankruptcy. According to Whitney (1993), club managers will, on average, overspend on talent that turns their team into a contender, i.e. they will over-invest in star players. The recent development of club finances in European soccer supports Whitney's hypothesis.

Késenne (2000*a*) develops a two-team model consisting of a large- and a small-market club and shows that a payroll cap, defined as a fixed percentage of league revenue divided by the number of teams, will improve competitive balance as well as the distribution of player salary within the league. Moreover, he shows that the profits of both the small- and the large-market club will increase.

The effect of salary caps on consumers (fans) has not been analyzed in the literature. This paper tries to fill the gap. We present a complete analysis of social welfare incorporating the effect of salary caps on clubs, players and fans. Using a game-theoretical model of a league consisting of both small- and large-market clubs, we show that salary caps will increase competitive balance and decrease the aggregate level of talent within the league. The resulting effect

[5] Going back to Rottenberg (1956) and Neale (1964).

[6] For a survey and discussion, see Szymanski (2003) and Borland and MacDonald (2003).

on social welfare is counter-intuitive and depends on the relative preference of fans for aggregate talent and for competitive balance. A salary cap that binds only for large-market clubs will increase social welfare if fans prefer aggregate talent despite the fact that the salary cap will result in lower aggregate talent. If fans prefer competitive balance, on the other hand, any binding salary cap will reduce social welfare.

The remainder of the paper is organized as follows. Section 2 outlines the basic model. In Section 3, we introduce salary caps into the model and distinguish different regimes depending on whether the salary cap is binding or not. Section 4 compares the aggregate salary payments, competitive balance and social welfare between the regimes. Section 5 concludes.

2 Model specification

The following model describes the impact of a salary cap on social welfare in a professional team sports league consisting of n (an even number) profit-maximizing clubs. The league generates total revenues according to a league demand function. The league revenue is then split among the clubs that differ with respect to their market share. We assume that there are two types of clubs, large-market clubs which receive a bigger share of league revenue and small-market clubs which receive a smaller share of league revenue. In order to maximize profits each club independently invests in playing talent where the supply of talent is assumed to be elastic. That is, the level of aggregate talent changes within the league depending on the salary payments. We regard the salary payment of each club as an investment in talent where the maximum amount that each club can invest in playing talent is defined by the salary cap.

League demand depends on the quality of the league q and is derived as follows.[7] We assume a continuum of fans who differ in their willingness to pay for a league with quality q. Every fan k has a certain preference for quality that is measured by θ_k. For simplicity, we assume that these preferences are uniformly distributed in $[0, 1]$, i.e. the measure of potential fans is one. Furthermore, we assume a constant marginal utility of quality and define the net-utility of fan θ_k as $\max\{\theta_k q - p, 0\}$. At price p the fan who is indifferent

[7] We derive league demand in exactly the same manner as Falconieri et al. (2004). Our approach, however, differs in an important aspect. For the sake of tractability we drop the contest theoretical part in the revenue function. Instead we use a slightly different quality function. The quality function q in Falconieri et al. (2004) is always increasing in own talent investments, i.e. $\frac{\partial q}{\partial x_i} > 0$, no matter how unbalanced the league becomes. In contrast, in our model, quality decreases if the league becomes too unbalanced (see also Runkel, 2006; Dietl and Lang, 2008).

to consumption of the league product or not is given by $\theta^* = \frac{p}{q}$.[8] Hence, the measure of fans who purchase at price p is $1 - \theta^* = \frac{q-p}{q}$. The league demand function is therefore given by $d(p,q) := 1 - \frac{p}{q}$. Note that league demand increases in quality, albeit with a decreasing rate, i.e. $\frac{\partial d}{\partial q} > 0$ and $\frac{\partial^2 d}{\partial q^2} < 0$. By normalizing all other costs (e.g. stadium and broadcasting costs) to zero, we see that league revenue is simply $LR = pd(p,q)$. Then, the league will choose the profit-maximizing price $p^* = \frac{q}{2}$.[9] Given this profit-maximizing price, league revenue depends solely on the quality of the league

$$LR = \frac{q}{4}. \tag{1}$$

Following the sports economic literature (e.g. Szymanski, 2003), we assume that league quality depends on the level of the competition, as well as the suspense associated with a close competition (competitive balance).[10] Moreover, we assume that the supply of talent is perfectly elastic and normalize the unit cost/price of talent to one. This means that talent investments of the clubs denoted by x are equal to their salary payments. In the following, we will use the two terms interchangeably.[11]

The level of the competition is measured by the aggregate talent within the n club league. We assume that the marginal effect of the salary payment (talent investment) x_i on the level of the competition T is positive but decreasing,[12]

$$T(x_1, .., x_n) = \alpha \sum_{j=1}^{n} x_j - \left(\sum_{j=1}^{n} x_j \right)^2. \tag{2}$$

[8]The price p can, for example, be interpreted as the subscription fee for TV coverage of the league.

[9]Note that the optimal price is increasing in quality, i.e. $\frac{\partial p^*}{\partial q} > 0$.

[10]According to Szymanski (2003) fan demand depends not only on the level of the competition and competitive balance but also on the 'likelihood of the home team's success'. Our results, however, are robust with respect to this specification: taking home team winning into consideration would result in an asymmetric quality function but would not alter our basic findings. For the sake of simplicity, we abstract from home team winning.

[11]Assuming a perfectly elastic supply of talent corresponds to a "worst-case scenario." In the other extreme case (fixed supply of talent), aggregate talent would remain constant in the league and a salary cap would improve competitive balance without having a negative effect on the level of the competition (aggregate talent). The introduction would therefore not result in a trade off between the level of the competition and competitive balance. As a consequence, a salary cap would always improve social welfare in the fixed-supply scenario.

[12]Note that our results are robust with respect to a linear talent function.

This is guaranteed in our model if $\frac{\partial T}{\partial x_i} > 0 \Leftrightarrow \sum_{j=1}^{n} x_j < \frac{\alpha}{2}$ and $\frac{\partial^2 T}{\partial x_i^2} < 0$ which will always be satisfied in equilibrium. Competitive balance CB is measured as minus the variance of salary payments[13]

$$CB(x_1, .., x_n) = -\frac{1}{n}\sum_{j=1}^{n}(x_j - \overline{x}_n)^2 \text{ with } \overline{x}_n = \frac{1}{n}\sum_{j=1}^{n} x_j. \tag{3}$$

Note that a lower variance of salary payments by the n clubs implies a closer competition and therefore a higher degree of competitive balance. League quality is now defined as

$$q(x_1, .., x_n) = \mu T(x_1, .., x_n) + (1-\mu)CB(x_1, .., x_n). \tag{4}$$

The parameter $\mu \in (0, 1)$ represents the relative weight that fans put on aggregate talent and competitive balance. Given aggregate salaries $\sum_{j=1, j\neq i}^{n} x_j$ of the other $(n-1)$ clubs, league quality increases in club i's salary payment x_i until a threshold value $x_i^*(\mu)$, i.e. $\frac{\partial q}{\partial x_i} > 0 \Leftrightarrow x_i < x_i^*(\mu)$. Since fans have at least some preference for competitive balance, excessive dominance by one club causes the quality to decrease.[14]

League revenues are split between the two types of clubs according to their market shares. For the sake of simplicity, we assume that half of the n clubs are large-market clubs which receive a bigger share of league revenue than the small-market clubs. Each of the large clubs receives a fraction $\frac{m_l}{n/2}$ of league revenues and each of the small clubs receives a fraction $\frac{m_s}{n/2}$ of league revenues, with

$$m_l > m_s \text{ and } m_l + m_s = 1. \tag{5}$$

We denote J_l and J_s as the set of large-market and small-market clubs, respectively, i.e. $J = \{1, .., n\} = J_l \cup J_s$.

The profit function $\Pi_i(x_1, .., x_n)$ of club $i \in J$ is given by revenue minus

[13]Obviously, there are different potential measures for competitive balance. We use the variance since this measure has the advantage of giving nice closed form solutions compared to other measures (e.g. coefficient of variation).

[14]Note that the threshold value $x_i^*(\mu)$ beyond which league quality decreases in club i's salary payments is an increasing function of the preference parameter μ because an increase in μ implies an increase in the preference for aggregate talent.

salary payments,

$$\Pi_i(x_1, .., x_n) = \frac{m_\delta}{2n}\left(\mu\alpha\sum_{j=1}^{n} x_j - \mu\left(\sum_{j=1}^{n} x_j\right)^2 - \frac{1-\mu}{n}\sum_{j=1}^{n}(x_j - \overline{x}_n)^2\right) - x_i, \tag{6}$$

with $\delta = l$ for $i \in J_l$ and $\delta = s$ for $i \in J_s$.

Social welfare is given by the sum of aggregate consumer (fan) surplus, aggregate club profit and aggregate player salaries. Aggregate consumer surplus CS corresponds to the integral of the demand function $d(p, q)$ from the equilibrium price $p^* = \frac{q}{2}$ to the maximal price $\overline{p} = q$ which fans are willing to pay for quality q,

$$CS = \int_{p^*}^{\overline{p}} d(p, q)dp = \int_{\frac{q}{2}}^{q} \frac{q-p}{q} dp = \frac{q}{8}. \tag{7}$$

Summing up aggregate consumer surplus, aggregate club profit and aggregate salary payments, social welfare is derived as

$$W(x_1, .., x_n) = \frac{3}{8} q(x_1, .., x_n). \tag{8}$$

Note that salary payments do not directly influence social welfare because salaries merely represent a transfer from clubs to players. As a consequence, social welfare depends only on the quality of the league.

3 Salary caps in a profit-maximizing league

Following Késenne (2000*a*), we introduce a salary cap into our model, which limits the total amount a club can spend on player salaries. The size of the salary cap, which is the same for each club, is based on the total league revenue in the previous season, divided by the number of clubs in the league. The salary cap *cap* is therefore exogenously given in the current season.

Clubs choose salary levels such that profits (6) are maximized subject to the salary cap constraint.[15] That is, salary payments x_i must not exceed the threshold *cap* given by the salary cap. The maximization problem for club

[15] For a discussion about the clubs' objective function see, e.g. Sloane (1971) and Késenne (2000*b*).

$i \in J$ is

$$\max_{x_i} \left\{ \frac{m_\delta}{2n} \left(\mu\alpha \sum_{j=1}^{n} x_j - \mu \left(\sum_{j=1}^{n} x_j \right)^2 - \frac{1-\mu}{n} \sum_{j=1}^{n} (x_j - \overline{x}_n)^2 \right) - x_i \right\}, \quad (9)$$

$$\text{subject to } 0 \leq x_i \leq cap,$$

with $\delta = l$ for $i \in J_l$ and $\delta = s$ for $i \in J_s$.

The corresponding first-order conditions are

$$\begin{aligned} \frac{m_\delta}{2n} \left(\mu \left(\alpha - 2\sum_{j=1}^{n} x_j \right) - \frac{2(1-\mu)}{n} \left(x_i - \frac{1}{n} \sum_{j=1}^{n} x_j \right) \right) - (1+\lambda_i) \leq 0, \\ x_i \left(\frac{m_\delta}{2n} \left(\mu \left(\alpha - 2\sum_{j=1}^{n} x_j \right) - \frac{2(1-\mu)}{n} \left(x_i - \frac{1}{n} \sum_{j=1}^{n} x_j \right) \right) - (1+\lambda_i) \right) = 0 \\ x_i - cap \leq 0 \\ \lambda_i (x_i - cap) = 0, \end{aligned} \quad (10)$$

where λ_i denotes the Lagrange multiplier for club $i \in J$ with $\delta = l$ for $i \in J_l$ and $\delta = s$ for $i \in J_s$.[16] To characterize the equilibrium, we have to distinguish different regimes depending on whether the salary cap is binding or not.

3.1 Regime A: salary cap is ineffective for all clubs

In this section, we assume that the salary cap is ineffective for all clubs, i.e. we consider the benchmark case that no (effective) salary cap exists.

In regime A, the equilibrium salary payments (talent investments) are computed from (10) as[17]

$$\begin{aligned} x_i^A &= \frac{\alpha}{2n} - \frac{m_l(1-\mu(1+n^2)) + m_s(1+\mu(n^2-1))}{2m_l m_s (1-\mu)\mu} =: x_l^A \ \forall i \in J_l, \\ x_j^A &= \frac{\alpha}{2n} - \frac{m_l(1+\mu(n^2-1)) + m_s(1-\mu(1+n^2))}{2m_l m_s (1-\mu)\mu} =: x_s^A \ \forall j \in J_s. \end{aligned} \quad (11)$$

For (11) to hold, in the following we restrict μ to $(\underline{\mu}, \overline{\mu})$.[18] The equilibrium

[16] It is easy to show that the second-order conditions for a maximum are satisfied.

[17] We denote the salary payments of club $i \in I$ in regime A with x_i^A. Analogous for regime B and C.

[18] For μ very close to zero or one the optimal choice for some clubs is zero. Since

salary payments show that all large-market (small-market) clubs choose the same salary level x_l^A (x_s^A). Note that without a binding salary cap the large-market clubs invest more in playing talent in equilibrium than the small-market clubs because the marginal revenue of talent investments is higher for these clubs. Thus, we are in regime A if in equilibrium the salary cap does not bind for the large-market clubs, i.e. if $cap \in I^A = [x_l^A, \infty)$.

In regime A, the aggregate level of salary payments $X^A = \sum_{j=1}^n x_j^A$ and competitive balance CB^A are given by

$$X^A = \frac{\alpha}{2} - \frac{n}{2\mu m_l m_s} \quad \text{and} \quad CB^A = -\left(\frac{n^2(m_l - m_s)}{2(1-\mu)m_l m_s}\right)^2. \tag{12}$$

Note that X^A (CB^A) is increasing (decreasing) in μ. That is, the higher the preference of fans for aggregate talent, the higher the aggregate salaries and the more unbalanced the league. The opposite holds if fans have a high preference for competitive balance.

Plugging the equilibrium salary payments (11) into equation (8) for social welfare yields the following level of total welfare in regime A

$$W^A = \frac{3}{32}\left(\mu\alpha^2 - \left(\frac{1}{\mu}\left(\frac{n}{m_l m_s}\right)^2 + \frac{1}{(1-\mu)}\left(\frac{n^2(m_l - m_s)}{m_l m_s}\right)^2\right)\right). \tag{13}$$

3.2 Regime B: salary cap is only effective for large-market clubs

In this section, we assume that the salary cap is only effective for the large-market clubs. That is, the salary cap constraint is only binding for club i with $i \in J_l$.

In regime B, the equilibrium salary payments (talent investments) are computed from (10) as

$$\begin{aligned} x_i^B &= cap =: x_l^B \ \forall i \in J_l, \\ x_j^B &= \frac{n(\alpha\mu m_s - 2n)}{m_s(1+\mu(n^2-1))} + cap\frac{1-\mu(n^2+1)}{1+\mu(n^2-1)} =: x_s^B \ \forall j \in J_s. \end{aligned} \tag{14}$$

we are not interested in a situation where clubs are not participating, we choose to restrict the range of μ to ensure positive equilibrium investments. Formally, we compute $(\underline{\mu}, \overline{\mu})$ as $\underline{\mu} = \frac{1}{2} - \frac{n^3(m_l-m_s)-n+((n^3(m_s-m_l)+n+\alpha m_l m_s)^2-4n\alpha m_l m_s)^{1/2}}{2\alpha m_l m_s}$ and $\overline{\mu} = \frac{1}{2} + \frac{n^3(m_s-m_l)+n+((n^3(m_s-m_l)+n+\alpha m_l m_s)^2-4n\alpha m_l m_s)^{1/2}}{2\alpha m_l m_s}$.

Thus, we are in regime B if $cap \in I^B = (cap', x_l^A)$ with $cap' := \frac{\alpha}{2n} - \frac{1}{\mu m_s}$. This condition guarantees that in equilibrium the small-market clubs invest less than cap. Otherwise the salary cap constraint would be binding for all clubs and regime C would be effective.

We now analyze how variations of the salary cap affect the clubs' optimal choice of salary payments. A more restrictive salary cap, i.e. a lower value of cap, induces the large-market clubs to decrease their salary payments in equilibrium, i.e. $\frac{\partial x_l^B}{\partial cap} > 0$. The effect on the small-market clubs' investment level is, however, ambiguous since

$$\frac{\partial x_s^B}{\partial cap} = \frac{1 - \mu(n^2+1)}{1 + \mu(n^2-1)} \begin{cases} > 0 \text{ if } \mu \in \left(\underline{\mu}, \frac{1}{n^2+1}\right), \\ = 0 \text{ if } \mu = \frac{1}{n^2+1}, \\ < 0 \text{ if } \mu \in \left(\frac{1}{n^2+1}, \overline{\mu}\right). \end{cases} \tag{15}$$

Hence, a more restrictive salary cap induces the small-market clubs to decrease their salary payments in equilibrium if $\mu \in \left(\underline{\mu}, \frac{1}{n^2+1}\right)$ and to increase their salary payments in equilibrium if $\mu \in \left(\frac{1}{n^2+1}, \overline{\mu}\right)$.[19] As a consequence, the higher the fans' preference for aggregate talent, the less talent is lost through a more restrictive salary cap.

What is the intuition for the result? The tightening of the salary cap has two effects on the investment incentives of the small-market clubs. On the one hand, a more restrictive cap lowers the salary payments by the large-market clubs and therefore enhances the incentive of the small clubs to pay higher salaries in order to "compensate" for the decrease in aggregate talent.[20] On the other hand, the incentive to improve competitive balance is weakened. If μ is relatively high, i.e. fans have a high preference for aggregate talent, then the first effect dominates the second effect and the small-market clubs increase their salary payments in equilibrium. If μ is relatively low, i.e. the fans have a high preference for competitive balance, then the incentive to improve competitive balance is lowered by the salary cap restriction so much that the small-market clubs will lower their salary payments in equilibrium. Finally, if $\mu = \frac{1}{n^2+1}$ then both effects balance each other out exactly.

The level of aggregate salary payments and competitive balance in regime

[19] Note that in equilibrium the small-market clubs never compensate for the reduction of talent by the large-market clubs owing to the salary constraint.

[20] Remember that quality is concave in aggregate talent.

B are given by

$$\begin{aligned} X^B(cap) =& \frac{n(1-\mu)}{1+\mu(n^2-1)} cap + \frac{n^2(\alpha\mu m_s - 2n)}{2m_s(1+\mu(n^2-1))} \quad \text{and} \\ CB^B(cap) =& -\left(\frac{n(2n+\mu m_s(2n\cdot cap - \alpha)}{2m_s(1+\mu(n^2-1))}\right)^2. \end{aligned} \tag{16}$$

Since $\frac{\partial x_l^B}{\partial cap} > \frac{\partial x_s^B}{\partial cap}$ a more restrictive salary cap will increase competitive balance and decrease aggregate salaries in regime B.

Social welfare in regime B is given by

$$W^B(cap) = \frac{3n(-n\mu(1-\mu)cap^2 + \alpha\mu(1-\mu)cap)}{8(1+\mu(n^2-1))} + \frac{3n(n\alpha^2\mu^2 m_s^2 - 4n^3)}{32m_s^2(1+\mu(n^2-1))}.$$

The highest attainable level of social welfare in regime B is obtained if the salary cap is fixed at

$$cap^B_{\max} = \frac{\alpha}{2n}. \tag{17}$$

In this case, social welfare is given by

$$W^B\left(\frac{\alpha}{2n}\right) = \frac{3\alpha^2\mu}{32} - \frac{3n^4}{8m_s^2(1+\mu(n^2-1))}. \tag{18}$$

We derive that if $\mu \in (\underline{\mu}, \mu']$ with

$$\mu' := \frac{1}{1+n^2(m_l - m_s)} \tag{19}$$

then social welfare will always decrease through the introduction of a binding salary cap.[21] We defer discussion of the implications to Section 4.

3.3 Regime C: salary cap is effective for all clubs

In this section, we assume that the salary cap is binding for the large-market and the small-market clubs. In this case, the equilibrium salary payments are simply given by

$$x_i^C = cap \text{ for all } i \in J. \tag{20}$$

We are in regime C if $cap \in I^C = (0, cap']$. Total salary payments $X^C(cap)$ are equal to $n \cdot cap$ and the competition is completely balanced with $CB^C = 0$.

[21]See Appendix A.1 for a derivation of conditions (17) and (19).

Social welfare in regime C is given by

$$W^C(cap) = \frac{3n}{8}\mu(-n \cdot cap^2 + \alpha cap). \tag{21}$$

4 Comparison of the regimes

By comparing the aggregate salary payments and competitive balance in regimes A, B and C, we derive the following proposition:

Proposition 1

(i) The level of competitive balance is decreasing in cap, i.e. $CB^C \geq CB^B(cap) \geq CB^A$.

(ii) The level of aggregate salaries is increasing in cap, i.e. $X^A \geq X^B(cap) \geq X^C(cap)$.

Proof. This result follows directly from (12), (16) and (20) and the definitions of I^k, $k \in \{A, B, C\}$. ■

This proposition shows that the introduction of a salary cap has the expected effect of increasing competitive balance and decreasing aggregate salaries.

By comparing social welfare in regimes A, B and C, we establish the following proposition:

Proposition 2

(i) If $\mu \in (\underline{\mu}, \mu']$*, i.e. fans prefer competitive balance, then an effective salary cap is always detrimental to social welfare.*

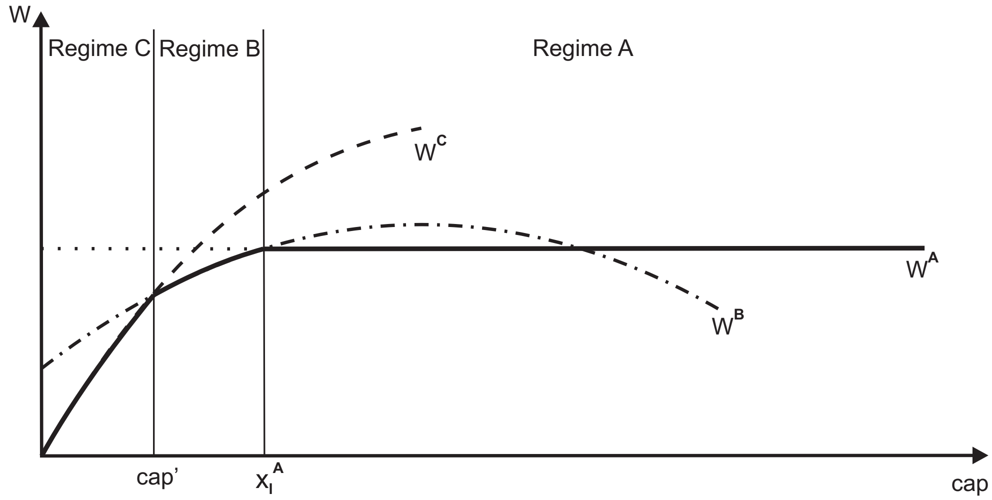
(ii) If $\mu \in (\mu', \overline{\mu})$*, i.e. fans prefer aggregate talent, then the highest attainable level of social welfare can be obtained in regime* B*.*

Proof. See Appendix A.2. ■

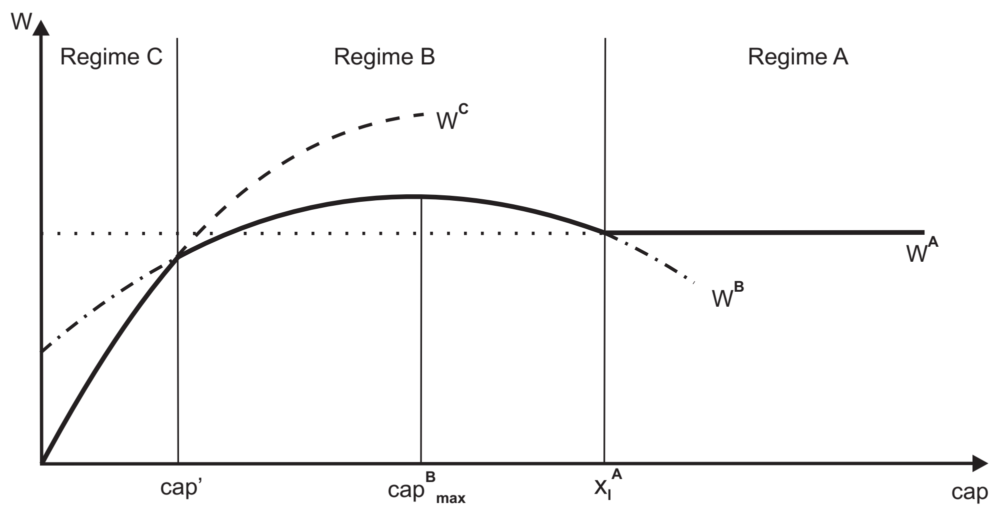
To see the intuition behind Proposition 2, consider Figure 1. The figure plots social welfare as a function of the salary cap for the case when fans prefer competitive balance (Figure 1a) and the case when fans prefer aggregate talent (Figure 1b). In both figures, the non-solid lines show the hypothetical levels of social welfare in regimes A to C, e.g. the level of social welfare in case the different regimes were effective. The level of social welfare in regime A is indicated by the dotted line (W^a). The dashed/dotted line (W^b) depicts social

Figure 1: Effect of Salary Caps on Social Welfare

a) Fans prefer competitive balance



b) Fans prefer aggregate talent



Notes: Hypothetical levels of social welfare in regimes $A - C$: W^A (dotted line), W^B (dashed/dotted line) and W^C (dashed line). The effective level of social welfare in regimes $A - C$ is indicated by the solid line.

welfare in case regime B is effective and the dashed line (W^c) social welfare in regime C. The solid line indicates the actual effective levels of social welfare in each regime. Imposing a stricter salary cap is characterized by a move along the x-axis from the right to the left hand side. Remember that regime A is only effective as long as the variable *cap* is above the investment level of the large-market clubs, i.e. $cap \geq x_l^A$. Whereas, regime B is active for $cap' < cap < x_l^A$ and regime C for $cap \leq cap'$. Also remember that a salary cap decreases aggregate talent in favour of a more even competition.

Figure 1a shows the case in which fans prefer competitive balance, i.e. $\mu \in (\underline{\mu}, \mu')$. Imposing a stricter salary cap (i.e. moving along the x-axis from the right to the left) does not affect social welfare in regime A since the salary cap constraint is not binding for any club. As soon as the salary cap crosses the threshold value x_l^A, the salary cap will bind for the large market clubs and we are in regime B. In this case, social welfare starts to decrease compared to regime A. This surprising result is because the unrestricted equilibrium in the case of a high preference for competitive balance is already characterized by a high level of competitive balance and a low level of aggregate talent. At these equilibrium values the marginal benefit of increased competitive balance through the salary cap is small, while the marginal loss owing to a decrease in aggregate talent is high (remember that for low μ less talent is lost through a more restrictive salary cap). In other words, there is no need to additionally increase competitive balance since the loss in aggregate talent outweighs the gains from a more even competition.

Social welfare decreases even faster in regime C, i.e. after crossing the threshold cap'. Imposing a stricter salary cap than cap' (implementing regime C) can never be optimal from a social point of view because the resulting loss in aggregate talent is not compensated by a positive effect on competitive balance, as the competition is already perfectly balanced.

This effect of the salary cap changes, however, as μ increases to $\mu > \mu'$, i.e. fans prefer aggregate talent. Figure 1b depicts this situation. Here, the unrestricted equilibrium in regime A is now characterized by a relatively high level of aggregate talent and a low level of competitive balance. Crossing the threshold value x_l^A, i.e introducing a binding salary cap for the large-market clubs will increase social welfare in regime B compared with regime A because the marginal benefit of increased competitive balance overcompensates for the marginal loss owing to a decrease in aggregate talent. This is true until the highest level of social welfare is attained at $cap = \frac{\alpha}{2n}$. Beyond that threshold, social welfare starts to decrease again, as the loss in talent cannot be overcompensated by the increase in competitive balance.

Note that if the fans' preference for aggregate talent increases beyond another threshold $\mu'' = \frac{3m_l+m_s}{3m_l+m_s+n^2(m_l-m_s)}$, i.e. $\mu \in (\mu'', \overline{\mu})$, then social welfare can also be higher in regime C than in regime A, although the highest level of social welfare is also attained in regime B.[22]

Moreover, it can be shown that if salary caps are beneficial for social welfare they also increase club profits.[23] Clubs will therefore never oppose salary caps which have a positive effect on social welfare. Caution is necessary, however, since there is a range of the preference parameter μ within which club profits increase and social welfare decreases through the introduction of a salary cap.[24]

5 Conclusion

Salary caps are employed within professional team sports leagues all over the world. As they are designed to limit salary payments, they could be interpreted as a collusive effort of club owners to control labor costs. On this assumption one could easily be inclined to predict that salary caps decrease social welfare. Using a game-theoretical model of a league consisting of small- and large-market clubs, we show that a salary cap may actually increase or decrease social welfare depending upon the fans' valuation of competitive balance and aggregate talent.

Our analysis shows that if the league is already very balanced (the case in which fans prefer competitive balance), a salary cap will reduce social welfare because it reduces the quality of the league by lowering the level of the competition. In this case the gains from additionally balancing the competition are low. In contrast, if the league suffers from an unequal distribution of talent (the case in which fans prefer aggregate talent), social welfare (and club profits) can be increased through a salary cap. The empirical research on the impact of within-season competitive balance suggests that the second case might be more realistic.[25] In any case, a binding cap will increase competitive balance and will help to keep salary costs under control. These results suggest that salary caps need not be a collusive effort but can be an important mechanism to increase social welfare (and club profits) within professional team sports leagues.

---

[22] See Figure 2 in Appendix A.3 for a graphical illustration of this situation.

[23] The analysis of club profits is similar to the analysis of social welfare.

[24] Formally: if the fans relative preference for aggregate talent is in the interval $(\widehat{\mu}', \mu')$ with $\widehat{\mu}' := \frac{n-4m_lm_s}{n^3(m_l-m_s)+n-4m_lm_s}$ then the introduction of a salary cap will be beneficial for the clubs and detrimental to social welfare.

[25] See e.g. Downward and Dawson (2000), Szymanski (2003) and Borland and MacDonald (2003). We are grateful to an anonymous referee for this point.

But why are salary caps not yet fully implemented in the European football leagues?

Although European club football has achieved economic and financial potential comparable to the US major leagues in the last decade, it has not followed the example of introducing salary cap mechanisms so far. Presumably, this reluctance is not due to the fact that the dangers of competitive imbalance and financial instability are unknown among the stakeholders of European football. The reasons for the past inactivity of European club football to introduce salary cap mechanisms are structural. In contrast to their American counterparts, football leagues in Europe are embedded in association structures. Every national football association governs a system of leagues, which is open through promotion and relegation from the amateur level to the top national division of professional football. At top of the national league pyramid, the European Football Association (UEFA), an association of national associations, organizes European club competitions like the Champions League and the UEFA Cup for the teams meeting certain sportive qualification criteria. The hermetic American major leagues operate independently of association structures and implemented salary caps as an integral part of a labour relations approach. The player's union and the owners represent the two sides of the relevant labour market and the state accepts the outcome of their bargaining written down in CBAs. Obviously, the labour relations approach employed by the hermetic American major leagues is not feasible within the European association-governed football pyramid. Football associations cannot be compared with the team owners in an American major league since associations do not represent one side of a labour market. Instead, associations are conceived as democratic governing bodies, which aim to integrate all important stakeholders of football in a certain geographic region including the players and, of course, the representatives of amateur football.

For historical and cultural reasons European states have left the regulation of sports to the sports governing bodies to a more or less substantial extent. This self-regulation of sports is seen as an important expression of European civil society. However, the scope for autonomous regulatory activity by the sports governing bodies is by no means unlimited. Recently the application of EU law has brought about a situation where the sports governing bodies have found it increasingly difficult to judge whether they are acting in accordance to EU law. The Bosman ruling of the EU Court of Justice is the most prominent case where a regulation issued by the football associations, the player transfer system, was found to violate EU law, in particular the principle of freedom of movement in the labour market. In this context it is a priori unclear if a salary cap mechanism in European football falls under the margin of discretion granted to the associations in order to perform their duties.

Due to the association based structure and the legal environment the decision-making processes concerning the introduction of salary caps will be much more complicated in the European context, as the interests of various stakeholders need to be properly balanced. Whether the European football leagues can cope with these challenges remains to be seen.

An interesting avenue for further research in this area is the welfare analysis of a "salary floor" in addition to a salary cap. A salary floor is the minimum amount of money a club must spend on team payroll. Note that the National Basketball Association (NBA), the National Hockey League (NHL) and the National Football League (NFL) already operate with such a salary floor. In the context of our model, a minimum team payroll increases aggregate talent and improves competitive balance. A salary floor could therefore help to improve the impact of salary caps on social welfare. A full formal treatment is left for future research.

A Appendix

A.1 Derivation of conditions (17) and (19)

We compute

$$\frac{\partial W^B(cap)}{\partial cap} = \frac{3}{8}\left(\frac{n(\alpha - 2n \cdot cap)(1-\mu)\mu}{1+\mu(n^2-1)}\right) > 0 \Leftrightarrow cap < cap^B_{\max} = \frac{\alpha}{2n}$$

The salary cap $cap^B_{\max}$ which induces the highest attainable level of social welfare need not, however, be within the interval of feasible salary caps $I^B = (cap', x^A_l)$ with $cap' = \frac{\alpha}{2n} - \frac{1}{\mu m_s}$ in regime B. We derive

$$x^A_l \leq cap^B_{\max} \Leftrightarrow \mu \leq \mu' := \frac{1}{1+n^2(m_l - m_s)} \in (\underline{\mu}, \overline{\mu}) \tag{22}$$

Hence, if $\mu \in (\underline{\mu}, \mu']$ then $cap^B_{\max} \notin I^B$ and a more restrictive salary cap, i.e. a lower variable cap, will decrease social welfare $W^B(cap)$ in regime B.

If, however, $\mu \in (\mu', \overline{\mu})$ then $cap^B_{\max}$ is in the interval of feasible salary caps I^B, i.e. $cap^B_{\max} \in I^B$. In this case the effect of a more restrictive salary cap on social welfare depends crucially on the size of the salary cap. Formally, we derive $\frac{\partial W^B(cap)}{\partial cap} > 0 \, \forall cap \in (cap', cap^B_{\max})$ and $\frac{\partial W^B(cap)}{\partial cap} < 0 \forall cap \in (cap^B_{\max}, x^A_l)$.

A.2 Proof of Proposition 2

This proof consists of three parts. In part (1) we compare regime A and B, in part (2) regime A and C and in part (3) regime B and C with respect to social welfare. Remember that regime A is only effective for $cap \geq x_l^A$, regime B for $cap' < cap < x_l^A$ and regime C for even tighter salary caps, $cap \leq cap'$.

(1) By comparing social welfare in regime A and B, we derive:

$$W^A \leq W^B(cap) \Leftrightarrow cap \in [cap_1^{AB}, cap_2^{AB}] \tag{23}$$

where

$$\begin{aligned} cap_1^{AB} &= \frac{\alpha}{2n} - \frac{m_l(1-\mu(1+n^2)+m_s(1+\mu(n^2-1))}{2m_l m_s(1-\mu)\mu} \text{ and} \\ cap_2^{AB} &= \frac{\alpha}{2n} + \frac{m_l(1-\mu(1+n^2)+m_s(1+\mu(n^2-1))}{2m_l m_s(1-\mu)\mu}. \end{aligned} \tag{24}$$

Note that cap_1^{AB} is exactly the equilibrium investment level of the large-market clubs in regime A, i.e. $cap_1^{AB} = x_l^A$.

We now analyze whether a salary cap from the interval $[cap_1^{AB}, cap_2^{AB}]$ for which social welfare is higher in regime B than in regime A is part of the interval I^B of feasible salary caps in regime B. We derive

$$cap_1^{AB} < cap_2^{AB} \Leftrightarrow \mu < \mu' := \frac{1}{1+n^2(m_l - m_s)} \in (\underline{\mu}, \overline{\mu}) \tag{25}$$

(1a) If $\mu \in (\underline{\mu}, \mu')$ then $\nexists\ cap' \in [cap_1^{AB}, cap_2^{AB}]$ such that $cap' \in I^B$. That is, we cannot find a salary cap out of the interval $[cap_1^{AB}, cap_2^{AB}]$ which is also included in the interval I^B of feasible salary caps for regime B. Hence,

$$W^A > W^B(cap)\ \forall cap \in (cap', cap_1^{AB}) = I^B \tag{26}$$

This shows that an effective salary cap is always detrimental to social welfare because social welfare is higher in regime A than in regime B.

(1b) If $\mu = \mu'$ then $W^A = W^B(cap) \Leftrightarrow cap = \frac{\alpha}{2n} = x_l^A$. Since $I^B = (cap', x_l^A)$ we also conclude that $W^A > W^B(cap)\ \forall cap \in I^B$.

(1c) If $\mu \in (\mu', \overline{\mu})$ then $cap_1^{AB} > cap_2^{AB}$. In this case, we have to analyze if cap_2^{AB} is in the interval of feasible salary caps I^B. We derive

$$cap_2^{AB} \leq cap' \Leftrightarrow \mu \geq \mu'' := \frac{3m_l + m_s}{3m_l + m_s + n^2(m_l - m_s)} \in (\mu', \overline{\mu}) \tag{27}$$

i) If $\mu \in (\mu', \mu'')$ then $cap' < cap_2^{AB}$ and thus the interval $[cap_2^{AB}, cap_1^{AB}]$ is a subset of the interval I^B. In this case the size of the salary cap determines whether social welfare is higher in regime A or B. More precisely, $W^A \geq W^B(cap)\ \forall cap \in (cap', cap_2^{AB}]$ and $W^A < W^B(cap)\ \forall cap \in (cap_2^{AB}, x_l^A)$

ii) If $\mu \in [\mu'', \overline{\mu})$ then $cap' \geq cap_2^{AB}$ and thus social welfare in regime B is higher than in regime A independent of the size of the salary cap, i.e. $W^A < W^B(cap)\ \forall cap \in I^B$.

Moreover, note that the highest level of social welfare is attained in regime B if the salary cap is fixed at $cap_{\max} = \frac{\alpha}{2n}$.

(2) By comparing social welfare in regime A and C, we derive:

$$W^A \leq W^C(cap) \Leftrightarrow cap \in [cap_1^{AC}, cap_2^{AC}] \tag{28}$$

where

$$\begin{aligned} cap_1^{AC} &= \frac{\alpha}{2n} - \frac{((n^2\mu m_l m_s)^2(1-\mu)(1-\mu+\mu n^2(m_l-m_s)^2))^{1/2}}{2(n\mu m_l m_s)^2(1-\mu)} \text{ and} \\ cap_2^{AC} &= \frac{\alpha}{2n} + \frac{((n^2\mu m_l m_s)^2(1-\mu)(1-\mu+\mu n^2(m_l-m_s)^2))^{1/2}}{2(n\mu m_l m_s)^2(1-\mu)}. \end{aligned} \tag{29}$$

We derive that $cap_1^{AC} < cap_2^{AC}$ and $cap_2^{AC} > cap'$, i.e. cap_2^{AC} is not in the interval of feasible salary caps I^C for regime C.

Analogously to (1), we analyze whether a salary cap from the interval $[cap_1^{AC}, cap_2^{AC}]$ for which social welfare is higher in regime C than in regime A is part of the interval $I^C = (0, \frac{\alpha}{2n} - \frac{1}{\mu m_s}]$ of feasible salary caps in regime C. We derive:

$$cap_1^{AC} \leq cap' \Leftrightarrow \mu \geq \mu'' := \frac{3m_l + m_s}{3m_l + m_s + n^2(m_l - m_s)} \tag{30}$$

(2a) If $\mu \in (\underline{\mu}, \mu'')$ then $cap_1^{AC} > cap'$. In this case cap_1^{AC} is not in the interval of feasible salary caps I^C for regime C and thus we derive that social welfare is higher in regime A than in regime C, i.e. $W^A > W^C(cap)\ \forall cap \in I^C$.

(2b) If $\mu \in [\mu'', \overline{\mu})$ then $cap_1^{AC} \leq cap'$. In this case the size of the salary cap determines whether social welfare is higher in regime A or C. More precisely, $W^A > W^C(cap)$ for all $cap \in (0, cap_1^{AC})$ and $W^A < W^C(cap)$ for all $cap \in (cap_1^{AC}, cap']$. Note that for $cap = cap'$ holds $W^A = W^C(cap)$.

Moreover, we derive that in regime C the highest level of social welfare would also be attained if the salary cap was fixed at $cap_{\max} = \frac{\alpha}{2n}$. However, this salary cap is never part of the interval I^C.

(3) By comparing social welfare in regime B and C, we derive:

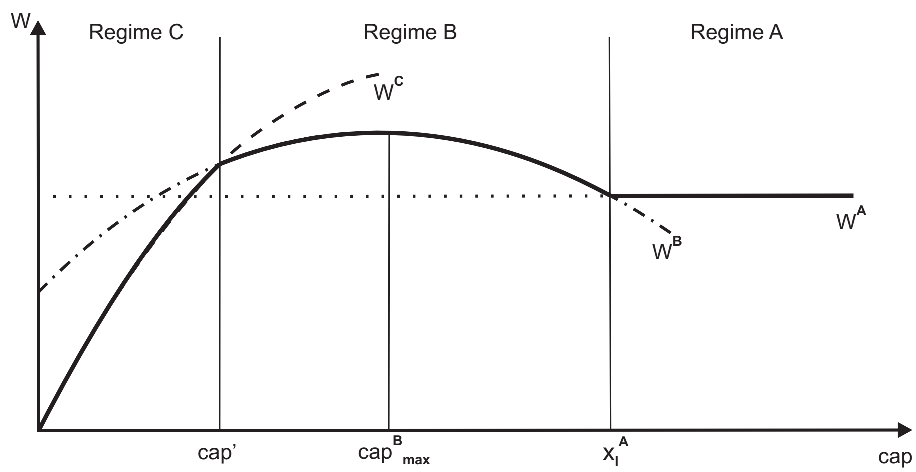
$$W^B(cap) \leq W^C(cap) \Leftrightarrow cap \in [cap_1^{BC}, cap_2^{BC}] \tag{31}$$

where

$$cap_1^{BC} = \frac{\alpha}{2n} - \frac{1}{\mu m_s} \text{ and } cap_2^{BC} = \frac{\alpha}{2n} + \frac{1}{\mu m_s} \tag{32}$$

Note that $cap_1^{BC} = cap'$. Moreover, regime C is only effective if $cap \in (0, cap']$. This directly implies that social welfare in regime C can never be higher than in regime B. As a consequence, implementing a sufficiently strict salary cap (i.e. $cap \leq cap'$) such that regime C is effective, will always decrease social welfare.

A.3 Figure 2: Effect of Salary Caps on Social Welfare for $\mu \in (\mu'', \overline{\mu})$



Notes: Hypothetical levels of social welfare in regimes A-C: W^A (dotted line), W^B (dashed/dotted line) and W^C (dashed line). The effective level of social welfare in regimes $A - C$ is indicated by the solid line.

References

Borland, J. and MacDonald, R. (2003), 'The demand for sports', *Oxford Review of Economic Policy* **19**, 478–502.

Dietl, H., Franck, E. and Lang, M. (2008), 'Overinvestment in team sports leagues: A contest theory model', *Scottish Journal of Political Economy* **55**(3), 353–368.

Dietl, H. and Lang, M. (2008), 'The effect of gate revenue-sharing on social welfare', *Contemporary Economic Policy* **26**, 448–459.

Downward, P. and Dawson, A. (2000), *The Economics of Professional Team Sports*, Routledge, London.

Falconieri, S., Palomino, F. and Sákovics, J. (2004), 'Collective versus individual sale of television rights in league sports', *Journal of the European Economic Association* **5**, 833–862.

Fort, R. and Quirk, J. (1995), 'Cross-subsidization, incentives, and outcomes in professional team sports leagues', *Journal of Economic Literature* **33**, 1265–1299.

Késenne, S. (2000*a*), 'The impact of salary caps in professional team sports', *Scottish Journal of Political Economy* **47**, 422–430.

Késenne, S. (2000*b*), 'Revenue sharing and competitive balance in professional team sports', *Journal of Sports Economics* **1**, 56–65.

Késenne, S. (2003), 'The salary cap proposal of the G-14 in european football', *European Sport Management Quarterly* **3**, 120–128.

Neale, W. (1964), 'The peculiar economics of professional sports: A contribution to the theory of the firm in sporting competition and in market competition', *Quarterly Journal of Economics* **78**, 1–14.

Rottenberg, S. (1956), 'The baseball players' labor market', *Journal of Political Economy* **64**, 242–258.

Runkel, M. (2006), 'Optimal contest design, closeness and the contest success function', *Public Choice* **129**, 217–231.

Sloane, P. (1971), 'The economics of professional football: The football club as a utility maximizer', *Scottish Journal of Political Economy* **17**, 121–146.

Staudohar, P. (1998), ‘Salary caps in professional team sports’, *Compensation and Working Conditions* pp. 3–11.

Staudohar, P. (2005), ‘The hockey lockout of 2004-05’, *Monthly Labor Review* **128**, 23–29.

Szymanski, S. (2003), ‘The economic design of sporting contests’, *Journal of Economic Literature* **41**, 1137–1187.

Vrooman, J. (1995), ‘A general theory of professional sports leagues’, *Southern Economic Journal* **61**, 971–990.

Vrooman, J. (2000), ‘The economics of American sports leagues’, *Scottish Journal of Political Economy* **47**, 364–398.

Whitney, J. (1993), ‘Bidding till bankrupt: Destructive competition in professional team sports’, *Economic Inquiry* **31**, 100–115.