OPTIMAL MARKET RESTRICTIONS IN PROFESSIONAL SPORTS[1]

William R. Dougan
Department of Economics
Clemson University

James M. Snyder, Jr.
Department of Political Science and Department of Economics
Massachusetts Institute of Technology

first version, March, 1994
this version, April, 2004

[1]We thank seminar participants at Clemson University, and participants at the workshops in Applications of Economics and Economics and Policy at the University of Chicago.

1. Competitive Restrictions in Professional Sports

Professional sports leagues in the United States coordinate the business decisions of their member firms to an extraordinary degree. Establishing a set of rules and a common schedule clearly require coordinating the output decisions of existing firms, but typical league activities go well beyond this minimal level. Leagues establish and award exclusive territorial franchises, and when new franchises are established their sale is performed by the league, which distributes the proceeds among its members. The movement of franchises from one area to another is also regulated by leagues, although to different degrees in different sports. Even more exceptional are the various restrictions on inter-team bidding for players that have been features of professional leagues for most of their existence. In this paper, we analyze the effects of entry restrictions and bidding restrictions to determine whether such practices, in addition to increasing the profits of existing teams, might confer any benefits upon the public at large.

While the establishment of regional franchises seems today to be virtually a defining feature of a sports league, the early days of professional baseball in the U.S. provide an example of organized play without exclusive territories. According to Scully (1989), by 1870 the National Association of Baseball Players counted over 400 amateur and professional teams among its members. In 1871, the professional teams withdrew to form the National Association of Professional Baseball Players. Over the next five years, 25 teams joined the professional association, although most were not financially viable. It was the formation of the National League in 1876 that introduced exclusive territorial franchises to professional baseball. Other innovations in league structure were the stipulation of a minimum city size for entry into the league and the payment of 30 percent of attendance revenue to visiting teams.

Within three years of its inception the National League developed a form of inter-team cooperation that would ultimately prove far more controversial than restricting entry by new teams. That innovation came to be known as the "reserve clause," which essentially granted each team exclusive rights to a number of players on its roster. Scully (p. 2) notes that "Club roster costs fell significantly after the assignment to the clubs of reservations rights to players. During the 1880s, for the first time, many clubs began to make a profit, and a semblance of stability emerged in the league."

Baseball's reserve clause lasted for 100 years before the advent of an essentially unrestricted market for players who have completed a minimum amount of time in the major leagues. Ironically, today's baseball player market is one of the least restricted of the major professional sports' labor markets (although the members of the American and National Leagues have been found guilty of less explicit but quite effective collusion in their bidding practices after the demise of the reserve clause). Until recently, National Football League players were allowed only limited free agency, and even today players with fewer than five years of league experience have only limited mobility. In addition, there are now restrictions on the total of each team's player salaries. Player transactions in the National Basketball Association are not directly restricted, but teams' total salaries are regulated. These restrictions on the salaries and movement of players across teams were negotiated in collective bargaining agreements between players' associations and the respective leagues.

Other restrictions on the market for players exist in all the major professional U.S. sports. One of these is a limit on the total number of players a team may carry on its roster. Such limits are taken for granted today, but were not in effect in the early days of professional baseball. For example, in 1909 Brooklyn had 61 reserved players on its roster while Washington had 29 (Scully, p. 4). Another seemingly universal practice is the use of so-called player drafts to assign to teams the initial rights to negotiate with players that have not been on the rosters of any other teams.

Practice similar to those common in professional sports would most likely be struck down as violations of the antitrust laws if they were practiced by other industries. They persist in the sports industry either by legislative exemption from the antitrust laws or, in the case of baseball, because the Supreme Court has ruled that those laws do not pertain. One justification for these exemptions has been that they are necessary to enable teams located in smaller markets to compete with teams from large markets in acquiring players, commonly known as the problem of competitive balance.

Virtually all economic analyses of competitive balance in sports leagues take the number of teams as fixed (El Hodiri and Quirk 1971, Quirk and El Hodiri 1974, Whitney 1993), and focus on the allocation of player talent across these teams. The presumption is that a chronically unequal distribution of talent is inherently undesirable. It is not clear, however,that aggregate welfare in these models is lower if teams from large cities win more games or championships than teams from smaller ones. After all, large cities have more fans deriving utility from the success of their teams. To see this in a different way, suppose all cities have equal populations, and differ only in the intensity of their fans' demand for winning. Would it be suboptimal for the city with the most devoted fans to win more often? At a minimum, competitive balance is not Pareto-superior to large-city domination, unless fans have a taste for balance *per se*.

In the paper we allow the number of teams to be determined endogenously. This allows us to examine the welfare consequences of certain deviations from competitive equilibria in a simple model in which optimal allocations are well defined. We find that if sports fans derive satisfaction simply from watching a talented home team,then the competitive equilibrium is efficient. However, if fans also care about their favorite team's overall performance relative to the rest of the league, then the unrestricted competitive equilibrium will generally not be efficient. Under plausible assumptions, too few teams will survive in equilibrium. Restrictions on inter-team competition similar to those actually observed can improve social welfare.[1]

As noted above, the key condition determining the optimality of competitive economic organization in professional sports is whether fans have a preference for a winning team over a non-winning team of equal ability.... Several empirical studies support the plausibility

[1]The number of teams is related to the issue of competitive balance. In our model all surviving teams may have identical talent, but some cities will not have teams. Thus, by determining the number of cities with teams, we in effect determine the size of the smallest city that will have a team. If we assume that in any sized league the largest cities will have teams, then the question "will Seattle and Milwaukee have teams?" is the same as the question "will there be 26 or 28 teams in the league?"

of this assumption. Noll (1974), Scully (1989), Whitney (1993), and Zimbalist (1992) all find that the attendance revenue of a major-league baseball team is positively related to its standing relative to the other teams in its league, and McCormick (1991) reports the same result for the teams of the National Basketball Association. This demand for winning teams generates several externalities that may render the purely competitive equilibrium suboptimal.

One of the welfare losses due to inter-team competition for players arises for the misallocation of labor that will occur if the supply of talent to teams is not perfectly inelastic (Canes, 1974). The marginal value of a player to any one term includes his contribution to that teams' standing relative to rival teams, but his social value does not include his effect on the relative positions of the various teams. For this reason, competitive bidding for the talent will cause players to be paid more than the value of the social marginal product, and the result is that society will have too many professional athletes relative to the optimum. In this paper we assume that the total pool of talent is exogenously given and that players have not alternative uses of their time, so our welfare analysis supplements that of Canes.

The second external effect of the demand for winning teams is a version of the common-pool problem. If the added competition from more teams does not sufficiently increase the satisfaction fans obtain from watching a winning team, then the revenue received by new entrants is only partially due to the fact that more fans have a local team to follow. Although the remainder is merely a transfer to new entrants from incumbents, because real costs are incurred to operate a team, the number of teams necessary to derive expected profits to zero is greater than the optimal number. We demonstrate in sections 4.1 and 4.2 the conditions under which the social optimum can be attained simply by restricting entry of new teams.

The third potential inefficiency of free competition in professional sports arises if there is a range of talent levels over which a team encounters increasing returns to talent. Whitney (1993) discusses the likelihood that this occurs when there is an asymmetric allocation of talent among teams. We explore this possibility in the context of a symmetric equilibrium in the market for player talent (section 4.3). We show that entry restrictions generally do not yield the social optimum because in equilibrium there are too few teams rather than too many.

In section 5 we show that, when leagues organize to maximize the joint profits of their members when confronting a competitive market for talent, there will always be too few teams in equilibrium. Inter league competition reduces but does not eliminate this inefficiency. We then show that certain restrictions on teams' bidding for talent can, at least in principle, induced profit-maximizing leagues to expand to the socially optimal number of teams.

Finally, in section 6 we recognize that the rules governing athletic competition "on the field" are themselves endogenous. We find that the desire of team owners to limit inter-team bidding for talent can lead them to adopt rules that increase the role of chance in determining the outcomes of games.

2. Competition for Talent in a Sports League

We study the following stylized model of a sports league. There are a potentially infinite number of *teams* that might compete in the league. Associated with each team i that competes is a level of *player talent* $Q_i \geq 0$. We follow Quirk and El Hodiri (1974) in defining the pool of available players in terms of a fixed quantity of homogeneous talent units. This does not mean that all players are identical, since each player may have a different amount of talent. We denote the total stock of talent as Q_T, so if there are n teams with positive talent levels, then $\sum_{i=1}^{n} Q_i \leq Q_T$. This assumption captures in a simple way the likelihood that the supply of very talented athletes is highly inelastic. It also ensures that the inefficiencies we find are not due to the misallocation of talent among professional sports teams.

Each team i generates two types of attendance, *normal attendance*, A_i, and *bonus attendance*, B_i. Normal attendance is a function of team talent, that is, $A_i = A(Q_i)$. This represents fans' preference over the absolute quality of play. It captures, for example, the fact that attendance at major-league baseball games is far greater than attendance at minor-league games. We assume the function A has the same shape for all teams, and A is twice continuously differentiable, strictly increasing and strictly concave in Q for all $Q \geq 0$. Also, we normalize by setting $A(0) = 0$.

Bonus attendance is a function of relative team performance. Specifically, there is a performance measure V and a bonus B such that team i receives a bonus attendance of BV_i if its performance level is V_i.[2] We assume $0 \leq V_i \leq 1$ for all teams i, and $V_i > 0$ only if $Q_i > 0$. Also, we normalize the V_i so $\sum_{i=1}^{n} V_i = 1$, where n is the number of teams with $Q_i > 0$.

There are at least two straightforward interpretations of V and B. First, we might imagine that fans care about where their team ranks in the league. In a league with n teams, we can assign performance levels $V^1, ..., V^k$ to the k teams with the highest winning percentages, where k may vary with n subject to $k \leq n$. These top k teams receive bonus attendances $BV^1, ..., BV^k$. For example, suppose $k = 2$, so only the top two terms receive bonus attendance, and suppose the first-place finisher receives twice the bonus of the second-place finisher (so $V^1 = 2/3$ and $V^2 = 1/3$). Then $B_i = 2B/3$ if team i finishes first, $B_i = B/3$ if it finishes second, and $B_i = 0$ otherwise.

Alternatively, we might imagine that fans care about their team's relative winning percentage. Then V_i gives the number of games won by team i as a fraction of the total number of games played by all teams in the league (this is also equal to team i's winning percentage divided by $n/2$), and B is the bonus attendance per fraction of games won. Clearly, the V_i may embody both interpretations simultaneously.

Implicit in this specification is the assumption that normal attendance is perfectly elastic with respect to ticket price whenever attendance is less than $A(Q_i)$ and perfectly inelastic at that point. Similarly, bonus attendance is perfectly elastic with respect to ticket price whenever bonus attendance is less than B, and perfectly inelastic at that point. We normalize by setting the reservation price of those who attend games equal to 1. This implies that

[2] More generally, B might be an arbitrary, increasing function of V. Treating such a general specification adds considerable notational and analytical complication, however, and tends to obscure rather than highlight our results.

social welfare is simply equal to total attendance (or revenue) net of social costs.[3]

From a social point of view, the talent pool Q_T is costless to use.[4] However, associated with each team is a fixed cost C. This reflects the costs of a stadium, administrative offices, and so on. Clearly, to maximize social welfare all of the talent must be used. Also, since A is strictly concave for all $Q \geq 0$, all teams that are assigned positive talent levels must be assigned the same talent level. Thus, the problem of maximizing social welfare can be written as

$$\max_{n,Q} \; nA(Q) + \sum_{i=1}^{n} BV_i - nC \quad \text{s.t.} \quad nQ = Q_T \tag{1}$$

or, since $\sum_{i=1}^{n} V_i = 1$,

$$\max_{Q} [A(Q) - C]Q + B \tag{2}$$

Differentiating (1) yields the first-order condition

$$A'(Q^*) = [A(Q^*) - C]/Q^* \tag{3}$$

That is, Q^* is chosen so that the marginal value of talent is equal to attendance revenue minus fixed costs, per unit of talent. Figure 1 shows Q^* graphically. Differentiating (1) twice with respect to Q, it is straightforward to verify that the second-order conditions are satisfied ($Q^* > 0$, and A is strictly concave for all $Q > 0$ by assumption). In fact, Q^* is the unique global optimum, since total surplus at Q^* is strictly positive ($A'(Q^*) > 0$, so $A(Q^*) > C$), and the only other possibility is the corner solution with $Q = 0$, which yields a total surplus of zero. The optimal number of teams is $n^* = Q_T/Q^*$. We assume that Q_T/Q^* is an integer.[5]

[3] There are two alternative assumptions that yield equivalent welfare implications. First, if teams can perfectly price discriminate, then gross revenue is equal to total consumer willingness to pay. Second, if there is a stadium capacity constraint that is binding in the relevant range of attendance, then social welfare is maximized by any price that fills the stadium.

[4] Alternatively, we might assume that the talent pool costs some amount D, but that D is independent of the amount of talent used provided the amount used is greater than zero.

[5] If Q_T/Q^* is not an integer, then the solution does not quite satisfy (2) but it is close. Specifically, the optimum n is either the largest integer less than Q_T/Q^*, or the smallest integer than Q_T/Q^*, and the optimal Q is Q_T/n.

Example. $A(Q) = Q^\alpha$, with $0 < \alpha < 1$. Then $A'(Q) = \alpha Q^{\alpha-1}$, so Q^* satisfies $\alpha(Q^*)^{\alpha-1} = [(Q^*)^\alpha - C]/Q^*$, or $Q^* = [C/(1-\alpha)]^{1/\alpha}$. Also, $n^* = Q_T[(1-\alpha)/C]^{1-\alpha}$. So for example, if $Q_T = 200$, $C = 1$, and $\alpha = 1/2$, then $Q^* = 4$ and $n^* = 50$.

3. Free Entry in the Absence of a Demand for Winning

Given our assumptions about demand and costs, when there is no bonus attendance a social optimum is achieved by a decentralized market mechanism.

Let w be the price of a unit of talent, *i.e.*, the wage. Assume that teams are price-takers in the talent market, paying w for each unit of talent they purchase, and that each team chooses its talent level to maximize profits. Teams are local monopolists, and therefore set ticket price equal to the fans' reservation price (which is 1). Then each team i solves

$$\max_{Q_i} \ \pi(Q_i) = \max_{Q_i} \ A(Q_i) - wQ_i - C \tag{4}$$

This yields the first-order condition $A'(Q_i) = w$, which clearly defines a unique profit-maximizing choice $Q_i = Q_O > 0$, provided that $A(Q_O) - wQ_O > C$. If $A(Q_O) - wQ_O < C$, then the unique profit-maximizing talent level is $Q_i = 0$, and if $A(Q_O) - wQ_O = C$ then either $Q_i = Q_O$ or $Q_i = 0$ is optimal.[6]

Define an *equilibrium* as a vector $(w, n, \mathbf{Q})$, where w is the wage, n is the number of teams competing in the league, and $\mathbf{Q} = (Q_1, ..., Q_n)$ is a vector of talent levels that satisfies: (i) for each team $i = 1, ..., n$, Q_i maximizes profits given w, (ii) all teams $i = 1, ..., n$ earn zero profits, and (iii) the market for talent clears, *i.e.*, $\sum_{i=1}^n Q_i = Q_T$. The proposition below shows that when Q_T/Q^* is an integer, the *unique* equilibrium produces a socially optimal allocation of talent.[7]

[6] Note that team i does not take the overall supply of talent into account when making its choice. That is, it does not worry about the fact that $\sum_{i=1}^n Q_i \leq Q_T$.

[7] To deal with cases where Q_T/Q^* is not an integer, we must revise the definition of equilibrium. The most natural way to do this is to replace (ii) with the following: (ii)' all teams $i = 1, ..., n$ earn nonnegative profits, and (ii)" if the league were to expand then all teams would earn negative profits, *i.e.*, for all $m > n$, if (i) and (iii) are satisfied then all teams $i = 1, ..., m$ earn negative profits. Then the unique equilibrium will be at the greatest integer less than Q_T/Q^*. We ignore these tedious complications, both here and below, by assuming that the simple equilibrium definitions result in whole number of teams.

Proposition 3.1. If $B = 0$ and Q_T/Q^* is an integer, then there is a unique equilibrium $(w_O, n_O, \mathbf{Q}_O)$, and this equilibrium satisfies $w_O = [A(Q^*) - C]/Q^*$, $n_O = Q_T/Q^*$, and $Q_{oi} = Q^*$ for all $i = 1, ..., n$.

Proof. All proofs are in an appendix.

Notice that the equilibrium wage can be written as $w_O = n[A(Q^*) - C]/Q_T$, which is just the total surplus divided by the total supply of talent. This is intuitive, since if teams are earning zero profits then all "rents" go to the players. Thus, maximizing social welfare maximizes total payments to players and also maximizes the wage.

4. The Welfare Implications of a Demand for Winning

When the bonus attendance is strictly positive, a decentralized market mechanism will generally not produce a social optimum. To develop the intuition for this, we analyze two cases. In the first case, all teams are assumed to have an equal chance of winning games (and the corresponding bonus attendance), regardless of the distribution of talent. Thus, a team's talent has no affect on its chances of winning. In the second case, each team's probability of winning depends on its talent level, as well as the talent levels of its opponents, teams with more talent naturally having higher probabilities of winning. While the first case is admittedly unrealistic, it provides an important intuition which is useful for understanding the more complicated second case.

4.1. Random Allocation of B

Suppose all teams with positive talent levels have an equal chance of winning games, regardless of their talent levels. Letting p_i denote the expected value of V_i, if $Q_i > 0$ for teams $i = 1..., n$, $p_i = 1/n$, regardless of $\mathbf{Q} = (Q_i, ..., Q_n)$. We call this the *random winner rule.*[8]

[8]The analysis in this section would also apply if gate-receipts were always shared equally by the home

As above, let w be the price of a unit of talent and assume that each team must pay w for each unit it purchases. Then, if $n-1$ teams other than team i choose positive talent levels, team i chooses its talent level Q_i to solve

$$\max_{Q_i} \; E\pi(Q_i) = B/n + A(Q_i) - wQ_i - C \tag{5}$$

This yields the first-order condition $A'(Q_i) = w$, which clearly defines a unique optimal choice $Q_i = Q_R > 0$, provided that $B/n + A'(Q_R) - wQ_R > C$. If $B/n + A'(Q_R) - wQ_R < C$, then the unique optimal choice for team i is $Q_i = 0$, and if $B/n + A'(Q_R) - wQ_R = C$, then either $Q_i = Q_R$ or $Q_i = 0$ is optimal.

As in the previous section, define an *equilibrium* as a vector $(w, n, \mathbf{Q})$ that satisfies: (i) for each team $i = 1, ..., n$, Q_i maximizes expected profits given w and n, (ii) all terms $i = 1, ..., n$ earn zero expected profits, and (iii) the market for talent clears, *i.e.*, $\sum_{i=1}^{n} Q_i = Q_T$. The next proposition shows that when $B > 0$, a socially optimal distribution of talent is *never* an equilibrium.

Proposition 4.1. If $B > 0$ and $(w_R, n_R, \mathbf{Q}_R)$ is an equilibrium under the random winner rule, then $n_R > n^*$ and $\mathbf{Q}_R$ satisfies $Q_{Ri} = Q^*$ for all $i = 1, ..., n_R$. (See Figure 2.)

Thus, with $B > 0$ there are *too many teams* in equilibrium, each with *too little talent*, relative to a social optimal.[9] The reason is simple: with $B > 0$, there is a negative externality associated with entry into the sport. When a new team enters it reduces the expected value of V_i for all previously existing teams (from $1/n$ to $1/(n+1)$, if n teams are competing before the new entrant). Potential teams do not take this into account in making their decisions to enter or not, so in equilibrium there are too many teams. This externality is the same as that associated with "common pool" problems.

How might a social optimum be achieved? Since the problem is a common pool type of

and visiting teams.

[9]Note that an equilibrium clearly exists provided that Q_T/Q_R is an integer. As above, we avoid tedious technical details by simply assuming that this is true.

externality, one obvious solution is to restrict entry to n^*, or to charge an appropriate entry fee. Therefore, we define a partial equilibrium with restricted entry as follows. For fixed n, $(w, \mathbf{Q})$ is a *restricted equilibrium* with n teams if: (i) for each team $i = 1, ..., n$, Q_i maximizes expected profits given w and n, (ii) all teams $i = 1, ..., n$ earn nonnegative expected profits, and (iii) the market for talent clears, *i.e.*, $\max_{n,Q}^{n} Q_i = Q_T$. Expected profits may be positive because entry is not permitted. Interestingly, when the social optimum is achieved at a restricted equilibrium, the wage w *falls*.

Proposition 4.2. Let $(w_R, n_R, \mathbf{Q}_R)$ be an equilibrium, and let $(w_R^*, \mathbf{Q}^*)$ be a restricted equilibrium with n^* teams. Then $w_R^* < w_R$.

It should be clear that $w_R^* = w_O$, that is, the restricted equilibrium wage is the same as the equilibrium wage for the case where $B = 0$. Also, expected profits for each team are $E\pi = B/n^* + A(Q^*) - w_R^* Q^* - C = B/n^*$, and total league profits are equal to the bonus B. That is, the teams, not the players, capture B. This is intuitive, since the bonus attendance B "occurs" regardless of the distribution of talent.

Another potential solution is to price the externality by giving all of the bonus B to the players, as part of a contingent compensation contract. That is, rather than simply pay a wage w, each team might offer to pay a wage w' and also pay any bonus attendance earned by the team to its players. It is straightforward to show that as long as players remain price-takers in the talent market, an equilibrium with this type of contact can produce a social optimum. A problem with the contract, however, is that it essentially makes the players the owners of the teams, and the price-taking assumption is therefore implausible. If instead the players act like owners and make joint decisions to maximize their expected compensation, the externality reappears.[10] Thus, entry restrictions or entry fees would appear to necessary in order to achieve a social optimum.

4.2. Teams Compete to Win B

[10] Possible player decisions include entry decisions, if groups of players collectively decide to form teams.

Now suppose that a team's relative success depends on its own talent level and on the talent levels of all other teams, but is stochastic. Specifically, assume that if there are n teams, then there are n functions $p_1, ..., p_n$ such that for each team i the expected value of V_i given the talent levels $\mathbf{Q} = (Q_1, ..., Q_n)$ is $p_i(\mathbf{Q})$. Assume the functions $p_1, ..., p_n$ have the following properties: (i) for each i, if $Q_i = 0$ and $Q_j > 0$ for some $j \neq i$, then $p_i(\mathbf{Q}) = 0$; (ii) each p_i is twice continuously differentiable, strictly increasing and strictly concave in Q_i; (iii) for each i, $0 \leq p_i(\mathbf{Q}) \leq 1$ for all $\mathbf{Q}$; (iv) $\sum_{i=1}^{n} p_i(\mathbf{Q}) = 1$ for all $\mathbf{Q}$; and (v) for all i and j, if $Q_i = Q_j$ then $p_i(\mathbf{Q}) = p_j(\mathbf{Q})$. Properties (i)-(ii) are self-explanatory, and properties (iii)-(iv) are necessary if $p_1, ..., p_n$ are to be interpreted as probabilities. Property (v) simply says that competition is symmetric across teams, so that if two teams choose the same quality level, then they have the same expected relative performance. Clearly, (iv) and (v) imply that if all teams choose the same talent levels then they will all have the same probabilities of winning, that is, if $\mathbf{Q}$ satisfies $Q_i = Q$ for all $i = 1, ..., n$, then $p_i(\mathbf{Q}) = 1/n$ for all $i = 1, ..., n$.

The assumption that each p_i is strictly concave in Q_i is not made because we believe it to be a good description of reality, but because it highlights the fact that competition for bonus attendance by itself is not sufficient to overcome the tendency for a competitively structured sports industry to provide too many rather than too few teams (this will be clear momentarily). In section 4.3, we study p_i with regions of increasing returns.

Each team is assumed to maximize expected profits. Since each team's probability of winning games depends on all teams' talent levels, in order to determine the choice of talent levels we must specify the teams' beliefs about each other's behavior. We assume "Nash conjectures," that is, each team i chooses its own talent level to maximize expected profits, taking all other teams' talent levels $\mathbf{Q}_{-i} = (Q_1, ..., Q_{i-1}, Q_{i+1}, ..., Q_n)$, as given.[11] Then,

[11] Thus, as in section 3, teams do not consider the constraint on the overall supply of talent when making their decisions. That is, they do not worry about the fact that $Q_1, ..., Q_n$ must satisfy $\sum_{i=1}^{n} Q_i \leq Q_T$. This assumption is less satisfactory than it was for the case when $B = 0$, because with $B > 0$ teams *must* make conjectures about all other teams' talent levels in order to solve their optimization problems, whereas with $B = 0$ they do not. One way to deal with this theoretically is to let teams' strategies be announcements of their *desired* talent levels, and assume that the total talent pool is then divided in proportion to the teams' announcements. That is, if $(R_1, ..., R_n)$ is the vector of announcements, each team i receives talent level $Q_i = R_i/(\sum_{i=1}^{n} R_j)$. Underlying this might be a "wage war" in which teams attempt to outbid each other for players, ultimately settling for the talent levels $Q_1, ..., Q_n$.

abusing notation somewhat, each i solves

$$\max_{Q_i} E\pi(Q_i; \mathbf{Q}_{-i}) = p_i(Q_i; \mathbf{Q}_{-i})B + A(Q_i) - wQ_i - C \tag{6}$$

This yields the first-order condition $\frac{\partial p_i}{\partial Q_i}(Q_i; \mathbf{Q}_{-i})B + A'(Q_i) = w$, which clearly defines a unique optimal choice $Q_i = Q_{Ci} > 0$, provided that $p_i(Q_{Ci}; \mathbf{Q}_{-i})B + A'(Q_{Ci}) - wQ_{Ci} > C$. If $p_i(Q_{Ci}; \mathbf{Q}_{-i})B + A'(Q_{Ci}) - wQ_{Ci} < C$, then the unique optimal choice for team i is $Q_i = 0$, and if $p_i(Q_{Ci}; \mathbf{Q}{-}i)B + A'(Q_{Ci}) - wQ_{Ci} = C$ then either $Q_i = Q_{Ci}$ or $Q_i = 0$ is optimal.

As in the previous sections, define an *equilibrium* as a vector $(w, n, \mathbf{Q})$ that satisfies: (i) for each team $i = 1, ..., n$, Q_i maximizes expected profits given w and $\mathbf{Q}_{-i}$, (ii) all teams $i = 1, ..., n$ earn zero expected profits, and (iii) the market for talent clears, *i.e.*, $\sum_{i=1}^{n} Q_i = Q_T$. Also, call an equilibrium *symmetric* if all teams with positive talent have the same talent level, *i.e.*, $Q_i = Q_T/n$ for all $i = 1, ..., n$.[12] The next proposition shows that when $B > 0$ and $p_1, ..., p_n$ satisfy properties (i)-(v), a socially optimal distribution of talent is *never* an equilibrium.

Proposition 4.3. If $(w_C, n_C, \mathbf{Q}_C)$ is an equilibrium with competition for bonus attendance, then $n_C > n^*$ and $\mathbf{Q}_C$ satisfies $Q_{Ci} < Q^*$ for all $i = 1, ..., n_C$.

As in the previous subsection, with $B > 0$ there are *too many teams* in equilibrium, each with *too little talent*, relative to a social optimum. Thus, it is again necessary to restrict entry or charge an entry fee in order to achieve an optimum.

Unlike the previous subsection, however, we cannot assert in general that the equilibrium wage is greater than the wage that would prevail in a socially optimal restricted equilibrium. In order to compare the two wages, we need to know more about the cross-partials of p_i. More specifically, if $(w_C^*, \mathbf{Q}^*)$ is a restricted equilibrium with n^* teams, then $w_C > w_C^*$ if $\frac{\partial p_i}{\partial Q_i}(Q_{Ci}; \mathbf{Q}_{C,-i}) - \frac{\partial p_i}{\partial Q_i}(Q_i^*; \mathbf{Q}_{-i}^*)$ is positive, or at least not too negative. In interesting special

[12] Unlike the previous section, in which only symmetric equilibria exist, with competition for bonus attendance non-symmetric equilibria might also exist.

cases, however, we can show that the equilibrium wage is too high. Suppose that for all $i = 1, ..., n$, p_i is defined by $p_i(Q_i; \mathbf{Q}_{-i}) = Q_i^b/(\sum_{j=1}^n Q_j^b)$, where $0 < b \leq 1$ ($b > 0$ is necessary to insure that p_i is increasing in Q_i, and $b \leq 1$ is necessary to insure that p_i is always concave in Q_i). While this functional form is not as general as some theorists would like, it has a number of desirable properties, and is used extensively in economics and political science, especially in the literatures on advertising, labor tournaments, rent seeking, voting and campaign finance.[13] It provides us with the following result.

Proposition 4.4. For all $i = 1, ..., n$, let p_i be defined by $p_i(Q_i; \mathbf{Q}{-i}) = Q_i^b/(\sum_{j=1}^n Q_j^b)$, with $0 < b \leq 1$. If $(w_C, n_C, \mathbf{Q}_C)$ is a symmetric equilibrium, and $(w_C^*, \mathbf{Q}^*)$ is a symmetric restricted equilibrium with n^* teams, then $w_C > w_C^*$.

As noted in the proof, in general $w_C^* > w_O$, where w_O is the wage that generates a social optimum when $B = 0$. Recall that w_O is also the wage that generates a social optimum with $B > 0$ at a restricted equilibrium under the random winner rule. Thus, compared to the random winner rule, when there is competition for the bonus the wage does not need to fall as far to achieve a social optimum. Also, expected profits of each team at a socially optimal restricted equilibrium are $E\pi = B/n^* + A(Q^*) - w_C^* Q^* - C = B[1/n^* - \frac{\partial p_i}{\partial Q_i}(Q^*; \mathbf{Q}_{-i}^*)Q^*]$, which is positive since p_i is strictly concave and $p_i(0; \mathbf{Q}_{-i}^*) = 0$. Total league profits are $n^* E\pi = B[1 - \frac{\partial p_i}{\partial Q_i}(Q^*; \mathbf{Q}_{-i}^*)Q_T]$. For the special case considered in proposition 4.4, this reduces to B/n^*. Thus, expected profits at a socially optimal restricted equilibrium are lower than expected profits under the random winner rule (recall that under the random winner rule, total expected profits are B).

Although free-entry equilibria are not socially optimal even when teams can compete for the bonus, competition does improve the situation somewhat relative to the case where the winning bonus attendance is random. Intuitively, competition for the bonus mitigates the externality associated with entry, because teams bid more for talent than when winning is

[13]See, *e.g.*, Friedman (1958), Tullock (1980), Rosen (1985), and Snyder (1989). Also, Elhodin and Quirk (19xx) use this in their analysis of sports leagues.

independent of talent, making entry less attractive.

Proposition 4.5. Let $(w_R, n_R, \mathbf{Q}_R)$ be an equilibrium under the random winner rule, and let $(w_C, n_C, \mathbf{Q}_C)$ be a symmetric equilibrium with competition for bonus attendance. Then $n_R > n_C$ and $Q_R < Q_C$.

4.3. The Likelihood of Ruinous Competition

The previous sections show that when there are continually diminishing returns to quality in the competition for the bonus attendance, then the competitive equilibrium results in too many teams, each with too little quality. However, correcting this market failure is straightforward, at least in principle, requiring only an entry restriction or entry fee. In this section, we show that it is sometimes necessary to do more than merely restrict entry in order to attain the social optimum.

In comparison with the assumption that each team's p_i function is strictly concave, it seems more plausible to assume that there is a range of talent levels over which a team's probability of winning bonus attendance exhibits increasing return. For example, suppose only the top of two teams in a league receive any bonus attendance. If the expected winning percentage of all teams in the league is 0.5, than a small increase in any one team's quality will have a significant impact on its probability of finishing in first or second place. However, a small increase in the quality of a team that expects to win only .33 of its games will have virtually no effect on its chances of coming in first or second.[14]

While we cannot say much generally about equilibria in the presence of such increasing returns, we can say quite a bit if we assume the functions p_i take on the functional form used in proposition 4.4. Thus, suppose that if n teams compete then $p_i(Q_i, \mathbf{Q}_{-i}) = Q_i^b / (\sum_{j=1}^n Q_j^b)$ for all $i = 1, ..., n$, where $b > 0$. In the proposition 4.4, we assumed that $b \leq 1$, which is

[14] Whitney (1993) gives another argument why the production function is likely to exhibit increasing returns. He shows that if the odds that a team wins any single game is proportional to its relative player talent (and games are independent events), then the probability that the team wins first place is convex in its relative talent, for many talent levels.

necessary and sufficient to guarantee that p_i is strictly concave in Q_i for all $Q_i > 0$ (provided that $Q_j > 0$ for at least one $j \neq i$). If $b > 1$, however, then p_i is first convex and then concave in Q_i, as suggested in the previous paragraph.

Proposition 4.6. Let p_i be defined by $p_i(Q_i; \mathbf{Q}_{-i}) = Q_i^b / \sum_{j=1}^n Q_j^b)$ for all $i = 1, ..., n$, with $b > 1$. Then any symmetric equilibrium $(w_C, n_C, \mathbf{Q}_C)$ satisfies the following: (i) if $b = n^*/(n^* - 1)$, then $n_C = n^*$ and $Q_C = Q^*$; (ii) if $b < n^*/(n^* - 1)$, then $n_C > n^*$ and $Q_C < Q^*$; and (iii) if $b > n^*/(n^* - 1)$, then $n_C < n^*$ and $Q_C > Q^*$.

Thus, it is possible, although not very likely, that competitive equilibrium with free entry will produce a social optimum.[15] It is more likely that there will be either too many teams each with too little talent, as in sections 3 and 4, or too few teams each with too much talent, relative to optimum.

Part (ii) of the proposition is an extension of the results in section 4. Since $n^*/(n^* - 1) > 1$, it shows that the results of section 4 sometimes hold even when the p_i are not everywhere concave. When $1 < b < n^*/(n^* - 1)$, p_i is not everywhere concave but the results of section 4 all hold.

Part (iii) describes a new case that demands further attention. To achieve a social optimum in this case one cannot simply impose an entry barrier or entry fee, since the problem is *too few teams*, not too many. In fact, one cannot even fix a wage and then allow teams to choose their talent levels as the next proposition shows.

Proposition 4.7. If $b > n^*/(n^* - 1)$, then there does not exist a wage w that will induce exactly n^* teams to optimally choose talent levels Q^*.

[15] We again omit the details of proving the existence of an equilibrium. In fact, it is straightforward to show that a symmetric equilibrium exists provided that Q_T/Q_C is an integer. Simply check that the second-order conditions for each team's profit maximization are satisfied at $(w_C, n_C, \mathbf{Q}_C)$. These conditions require that $n_C < 2b/(b-1)$, so for $b > 1$ they impose an upper bound on the number of teams that compete in equilibrium. For b large enough, the number of competing teams must be two.

To achieve a social optimum in this case requires either a *subsidy* to entry, or explicit restrictions on teams' talent levels, or a *tax* on player salaries with the proceeds refunded to the teams in a lump sum fashion. It is straightforward to calculate the optimal entry subsidy or talent tax (the tax rate given in the proposition is optimal under the assumption that teams take total tax revenue as fixed).

Proposition 4.8. If $b > n^*/(n^*-1)$, then under either of the following policies there exists a symmetric equilibrium $(w_C, n_C, \mathbf{Q}_C)$ with $n_C = n^*$ and $Q_{Ci} = Q^*$ for all i: (i) a subsidy of $S = B[b(n^*-1) - n^*]/(n^*)^2$ to each team that competes; or (ii) a tax of $t = B[b(n^*-1) - n^*]/n^* Q_T$ for each unit of talent purchased by a team, with tax proceeds distributed equally among the competing teams.

5. Equilibrium When League Members Maximize Joint Profits

The analysis of the previous sections assumed that although individual teams sought to maximize their profits, they did not exploit the league structure to prevent themselves from driving expected profits to zero. In this section, we assume that leagues act to minimize the total profits of their member teams, and explore the implications of this assumption for league size and organizational structure. If we assume that each league has a small number of "founders" with exclusive rights to sell franchises, this assumption makes sense even when the number of teams in a league is itself a choice variable. The founders can be expected to choose league size in order to maximize their combined profits from the operation of their own teams and the sale of new franchises, which will lead them to choose the number of teams that maximizes total league profits.

First, suppose there is one league, which can restrict entry and thereby control the number of teams that compete. Assume also that once the number of teams in the league is determined, the teams then compete freely for talent (the league is not assumed initially to have the power to prohibit such competition). Also assume the probability each team i

wins the bonus is given by the function $p_i(Q_i, \mathbf{Q}_{-i}) = Q_i^b/(\sum_{j=1}^n Q_j^b)$, as above.[16] Finally, supposed that the league restricts entry in order to maximize the joint profits of its member teams, assuming that the teams' subsequent competition for talent will produce a symmetric restricted equilibrium (recall the definition of a restricted equilibrium in section 3). Then, as the next proposition shows, the league will have too few teams, each with too much talent, relative to the social optimum.

Proposition 5.1. Let n_1 be the number of teams that maximizes the joint profits of the league's member teams as the symmetric restricted equilibrium, and let Q_1 be the amount of talent on each team at the equilibrium. Then $n_1 < n^*$ and $Q_1 > Q^*$.

The intuition behind this result is straightforward. Since the league is interested in joint profits, players' salaries are viewed as costs in deciding on the number to teams. Thus, the league wants to keep wages low. Since competition among the teams will ensure that players are paid their marginal project, this means keeping the marginal product of talent low. This is achieved by reducing the number of teams and increasing the talent on each team, since the marginal product of talent on each team is decreasing in the team's total talent level.

One possible way to improve the situation is to allow the formation of new leagues that compete with one another for talent. However, as we now show, a social optimum cannot be achieved simply by adding leagues.[17] Supposed there are L leagues, and suppose that bonus attendance is divided equally among the leagues, so each league has its own bonus of B/L (so total bonus attendance is fixed). Also, suppose each league restricts entry in order to maximize the joint profits of its member teams, assuming that the subsequent competition for talent will result in a restricted symmetric equilibrium in each league. Also, assume that the leagues play Nash against one another, simultaneously choosing their sizes. Then we have the following.

[16]Results similar to those below also hold for the random winner rule.

[17]In fact, it is possible that adding leagues can make the situation worse. We do not prove this result here in the interests of space.

Proposition 5.2. Suppose L leagues compete with one another for talent, each league choosing the number of teams that maximizes its members' joint profits at a symmetric restricted equilibrium. Let n_L be the number of teams in each league at the symmetric restricted equilibrium in which all leagues are the same size, and let Q_L be the corresponding talent level of each team. Then $Ln_L < n^*$ and $Q_L > Q^*$.

There are two ways to achieve a socially optimal allocation of talent with one league. One is to allow the league to tax player salaries, refunding the tax revenues to the teams in a lump-sum fashion. This effectively allows the league to drive players' wages to zero, at which point the league maximizes joint profits by maximizing joint revenue, and thereby maximizes social welfare.

Proposition 5.3. Suppose there is one league, which can restrict entry and can also impose a tax of t per unit of talent each team buys, with the tax revenues distributed across the leagues's teams in a lump-sum fashion. Let n_τ and t_τ be the number of teams and tax rate, respectively, that maximize the league's joint profits at a symmetric restricted equilibrium, and let Q_τ be the amount of talent on each team at the equilibrium. Then $n_\tau = n^*$, and $Q_\tau = Q^*$. Also, the equilibrium wage is $w_\tau = 0$.

The equilibrium wage is zero only because of our assumption that players' opportunity wages are zero. If players have a positive opportunity wage of, say, $\underline{w}$, then the equilibrium wage will be $\underline{w}$. However, the number of teams, and the talent on each team, will still be socially optimal. Of course, if equity in the compensation of owners and players is a consideration, this solution may be viewed as less than fully satisfactory, since all rents generated by the leagues accrue to the owners.

A second solution is to impose a ceiling on the total wage bill each team is allowed to pay that depends on the team's total expected revenues net of fixed costs. That is, suppose there exists $\theta \in (0, 1)$ (perhaps chosen by a social planner, perhaps the results of a bargain collectively struck by players and team owners), such that no team i is allowed to pay more than $\theta[B/n + (1/n)\sum_{i=1}^{n} A(Q_i) - C]$ in total salaries to its players. If this ceiling on salaries is low enough, then a single league that can restrict entry will admit the socially optimal number of teams.

Proposition 5.4. Suppose that for all teams i, Q_i must satisfy $wQ_i \leq \theta[B/n+(1/n)\sum_{i=1}^{n} A(Q_i) - C]$. Let n_θ be the number of teams that maximizes the league's joint profits at a symmetric restricted equilibrium, and let Q_θ be the amount of talent on each team at the equilibrium. If $\theta < 1 - \pi_1(n_1)/S^*$, where $\pi_1(n_1)$ is the maximized value of total league profits in the absence of a ceiling on salaries (defined in Proposition 5.1), and S^* is the maximized value of total surplus, then $n_\theta = n^*$ and $Q_\theta = Q^*$.

The upper bound on θ exists because the league can always achieve profits greater than or equal to the profits associated with limited entry and free competition for talent (the arrangement in proposition 5.1). The lower bound on θ is zero, since we have assumed that players' reservation wages are zero. If players have positive reservation wages, then the lower bound on θ is greater than zero.

Each of these restrictions on inter-team competition for player talent leads teams and leagues toward the social optimum because they render the problem of maximizing joint team profits identical to the social welfare maximization problem. Clearly, any other type of restriction that equates profit maximization with surplus maximization will serve equally well. We have explicitly examined these two restrictions because they resemble practices that have actually been observed in the sports industry.

The profit-maximizing league-imposed tax induces each team to act as a monopsonist in the market for talent. Thus, a properly designed tax-cum-rebate would mimic the system followed by major league baseball until the demise of the reserve clause in 1975. The upper bound on salaries resembles the "salary cap" currently in effect in the National Basketball

Association, and the conditional salary cap in the National Football League.[18] Under the system, individual players and teams participate in a market for free agents, but each team's total payroll is limited. Under the rule we consider, each team's total payroll is predetermined and each player receives θ times the average product of talent for each unit of talent that player possesses.[19] Of course, we do not know whether the mix of restrictions practiced in these leagues have actually served to increase total welfare.

6. Choosing the Rules of the Game

In sections 4 and 5, we take the parameter b in $p_i(Q_i, \mathbf{Q}_{-i}) = Q_i^b/(\sum_{j=1}^n Q_j^b)$ as exogenous. It is possible, however, that b can be varied by changing the rules that govern league play. The parameter b essentially reflects the extent to which outcomes are uncertain rather than deterministic. If b is close to zero, then outcomes are virtually random, while as b increases they become increasingly deterministic; in the limit, a team with slightly more talent wins with certainty.

Neither of these extremes seems likely to maximize fans' enjoyment from seeing their team win. If outcomes were totally random, victory would be of particular significance by itself. On the other hand, if there were no uncertainty over outcomes, the element of suspense would be absent. (The first case is analogous to watching a game or series of games in which, regardless of what happened on the field, the winner is determined by a coin toss. The second is analogous to watching videotapes of games whose outcomes are already known.) This is not to say that games with these features would be of no spectator interest. Fans will still want to see highly talented athletes perform – just as they go to the opera or theater – but

[18]The salary caps in the NBA are somewhat flexible. Specifically, when one of a team's free agents has been bid away by another team, the team losing the free agent may exceed its salary caps in the year the free agent moves.

[19]A different sort of league policy that is sometimes proposed is revenue sharing among teams. Full revenue sharing would clearly eliminate the inefficiencies arising from competition for bonus attendance, since it offers each team the same expected return as the case in which the bonus is randomly assigned. Thus, a social optimum *could* be achieved through revenue sharing together with the appropriate entry restriction, a join profit-maximizing league would instead restrict itself to the minimum feasible number of teams, in order to minimize costs.

this is captured in the normal attendance function $A(Q)$. What it does suggest, however, is that the bonus attendance, B, may itself depend on the value of b for a particular sport.

Since leagues determine the rules of play for their members, it is natural to extend the model by allowing them to choose the value of b. As the next proposition shows, a joint-profit-maximizing league will specify a set of rules that introduce too much randomness relative to the social optimum, *i.e.*, the league will choose too low a value of b.

Proposition 6.1. Suppose B is a strictly concave, twice differentiable function of b, which achieves its maximum at $b^* > 0$. Let (b_b, n_b) be the degree of determinism and number of teams that maximize the joint profits of league's member teams at the symmetric restricted equilibrium. Then b^* is the socially optimal degree of determinism, and $n_b < n^*$ and $b_b < b^*$.

The intuition for this result is straightforward: if the degree of determinism is less than b^*, then increasing the degree of determinism increases the size of the bonus, but it also increases inter-team competition for talent. This drives up wages, and therefore reduces teams' profits. The league therefore chooses $b < b^*$. From a social point of view, however, wages are not costs but merely transfers from team owners to players, so the socially optimal value of b is b^*.

7. Conclusion

Free entry into unrestricted competition among professional sports teams is unlikely to sustain the socially optimal number of teams when the revenue of each team depends in part on its performance relative to its competitors. When there are diminishing returns to team quality, the unrestricted symmetric equilibrium involves too many teams, each with too little talent, relative to the optimum. This can be interpreted either as referring to the (endogeneous) rooster size of each professional team, as was the case in the early years of baseball, or as referring to the total number of players under contract to an organization at any level of play, as in the case of baseball and hockey teams' minor league systems today.

On the other hand, when there is a range of talent levels that exhibit increasing returns, as seems plausible in most sports, freely competing teams will tend to stockpile too much talent. The equilibrium number of teams will be below the social optimum, so that no form of simple entry restriction can improve the situation. In this case, certain types of restrictions on inter-team competition in the market for players can induce a welfare-increasing allocation of talent among incumbent and newly entering teams. In order for those restrictions to induce a social optimum, they must render profit-maximizing choices by teams and leagues to be identical to the corresponding welfare-maximizing choices. The most straightforward such rule is a predetermined upper bound on the wage bill for each team that is calculated as a constant share of total team revenue.

Appendix

Proof of Proposition 3.1. If team i maximizes profits by choosing $Q_i = Q_O > 0$ and also earns zero profits, then $\pi(Q_O) = A(Q_O) - A'(Q_O)Q_O - C = 0$. By (2) above, this implies that $Q_O = Q^*$. Thus, $(w_O, n_O, \mathbf{Q}_O)$ is an equilibrium only if $Q_{oi} = Q^*$ for all $i = 1, ..., n$. Also, $w_O = [A(Q^*) - C]/Q^*$. Finally, by part (iii) of the definition of an equilibrium, $n_O Q^* = Q_T$, so $(w_O, n_O, \mathbf{Q}_O)$ is an equilibrium only if $n_O = Q_T/Q^*$. Thus, the only possible equilibrium has $w_O = [A(Q^*) - C]/Q^*$, $n_O = Q_T/Q^*$, and $Q_{oi} = Q^*$ for all $i = 1, ..., n$, and it is straightforward to verify that this is in fact an equilibrium.

Proof of Proposition 4.1. If each team $i = 1, ..., n_R$ maximizes expected profits by choosing $Q_i = Q_R > 0$ and also earns zero expected profits, then $E\pi(Q_R) = B/n_R + A(Q_R) - A'(Q_R)Q_R - C = 0$, or $A'(Q_R) = [A(Q_R) - C]/Q_R - B/(n_R Q_R)$. Thus, $A'(Q_R) > [A(Q_R) - C]/Q_R$, which means that $Q_i < Q^*$. Thus, $(w_R, n_R, \mathbf{Q}_R)$ is an equilibrium only if $Q_{Ri} = Q_R < Q^*$ for all $i = 1, ..., n$. By part (iii) of the definition of an equilibrium, $n_R Q_R = Q_T$, so (W_R, n_R, Q_R) is an equilibrium only if $n_R > Q_T/Q^* = n^*$.

Proof of Proposition 4.2. The first-order condition for each team's expected profit maximization problem implies $w_R^* = A'(Q^*)$. From proposition 4.1 above, if $(w_R, n_R, \mathbf{Q}_R)$ is an equilibrium, then $w_R = A'(Q_R)$ and $Q_R < Q^*$. And A is strictly concave, so $A'(Q_R) > A'(Q^*)$, and thus $w_R > w_R^*$.

Proof of Proposition 4.3. If team i maximizes expected profits by choosing $Q_{Ci} > 0$ and also earns zero expected profits, then $E\pi(Q_{Ci}; \mathbf{Q}_{C,-i}) = p_i(Q_{Ci}; \mathbf{Q}_{C,-i})B + A(Q_{Ci}) - [\frac{\partial p_i}{\partial Q_i}(Q_{Ci}; \mathbf{Q}_{C,-i})B + A'(Q_{Ci})]Q_{Ci} - C = 0$. Rearranging yields $A'(Q_{Ci}) - [A(Q_{Ci}) - C]/Q_{Ci} = [p_i(Q_{Ci}; \mathbf{Q}_{C,-i}) - \frac{\partial p_i}{\partial Q_i}(Q_{Ci}; \mathbf{Q}_{C,-i})]B$. Since p_i is strictly increasing and strictly concave in Q_i and $p_i(\mathbf{Q}) = 0$ when $Q_i = 0$, $\frac{\partial p_i}{\partial Q_i}(Q_{Ci}; \mathbf{Q}_{C,-i}) < p_i(Q_{Ci}; \mathbf{Q}_{C,-i})/Q_{Ci}$. Thus, $A'(Q_{Ci}) - [A(Q_{Ci}) - C]/Q_{Ci} > 0$, which means that $Q_{Ci} < Q^*$. Thus, $(w_C, n_C, \mathbf{Q}_C)$ is an equilibrium only if $Q_{Ci} < Q^*$ for all $i = 1, ..., n$. By part (iii) of the definition of an equilibrium,

$\sum_{i=1}^{n} Q_{Ci} = Q_T$; and $\sum_{i=1}^{n} Q_{Ci} < n_C Q^*$, so $(w_C, n_C, \mathbf{Q}_C)$ is an equilibrium on if $n_C > Q_T/Q^* = n^*$.

Proof of Proposition 4.4. Given the functional form of p_i, it is straightforward to show that $\frac{\partial p_i}{\partial Q_i}(Q_i; \mathbf{Q}_{-i}) = bp_i(Q_i; \mathbf{Q})[1-p_i(Q_i; \mathbf{Q}_i)]/Q_i$ for all $Q_i > 0$ (differentiate with respect to Q_i, divide the result by Q_i, and substitute). Thus, the first-order condition for each team's expected profit maximization problem implies that at a symmetric restricted equilibrium $W_C^* = Bb(n^*-1)/((n^*)^2 Q^*) + A'(Q^*) = Bb(n^*-1)/(n^* Q_T) + A'(Q^*)$. (Note that if w_C^* is the wage that generates a social optimum under the random winner rule, then in general $w_C^* > w_O$). From proposition 4.3 above, if $(w_C, n_C, \mathbf{Q}_C)$ is a symmetric equilibrium, then $w_C = Bb(n_C - 1)/(n_C^2 Q_C) + A'(Q_C) = Bb(n_C - 1)/(n_C Q_T) + A'(Q_C)$ and $n_C > n^*$. Since A is strictly concave in Q_i, $A'(Q_C) > A'(Q^*)$. Also, $n_C > n^*$ implies that $(n_C - 1)/n_C > (n^*-1)/n^*$. Thus, $w_C > w_C^*$.

Proof of Proposition 4.5. By the definition of equilibrium, $A'(Q_R) - [A(Q_R) - C]/Q_R = B/Q_T$ and $A'(Q_C) - [A(Q_C) - C]/Q_C = B/Q_T - B[\frac{\partial p_i}{\partial Q_i}(Q_C; \mathbf{Q}_{C,-i})$ (see the proofs of propositions 4.1 and 4.3, respectively), so $A'(Q_R) - [A(Q_R) - C]/Q_R > A'(Q_C) - [A(Q_C) - C]/Q_C$. Also, $Q_R < Q^*$ and $Q_C < Q^*$, where Q^* solves $[A(Q^*) - C]/Q^* = A'(Q^*)$. Since A is strictly concave and $C > 0$, $A'(Q) - [A(Q) - C]/Q$ is strictly decreasing in Q for all $Q < Q^*$ (this is easily verified by differentiation), so $Q_R < Q_C$. And, since $n_R Q_R = n_C Q_C = Q_T$, $n_R > n_C$.

Proof of Proposition 4.6. Recall from the proposition 4.4 that $\frac{\partial p_i}{\partial Q_i}(Q_i; \mathbf{Q}_{-i}) = bp_i(Q_i; \mathbf{Q}_{-i})[1-p_i(Q_i; \mathbf{Q}_{-i})]/Q_i$ for all $Q_i > 0$. Thus, the first-order condition for each team's expected profit maximization problem implies that at a symmetric equilibrium, $Bb(n_C - 1)/(n_C^2 Q_C) + A'(Q_C) = w_C$. The zero-profit condition is $B/n_C + A(Q_C) - C - w_C Q_C = 0$, and market-clearing requires that $w_C Q_C = Q_T$, so combining these with the first-order conditions and eliminating w_C and n_C yields $A'(Q_C) - [A(Q_C) - C]/Q_C = B(1-b)/Q_T + Q_C Bb/(Q_T^2)$. The left-hand side of this equation is zero at Q^*, positive and strictly decreasing in Q_C for all

$Q_C < Q^*$, and negative for all $Q_C > Q^*$. The right-hand side is negative at $Q_C = 0$ (since $b > 1$), and strictly increasing in Q_C. Thus, substituting, if $b = n^*/(n^*-1)$ then the equation is uniquely satisfied at $Q_C = Q^*$. If $b < n^*/(n^*-1)$ then the equation is uniquely satisfied at some $Q_C < Q^*$, and if $b > n^*/(n^*-1)$ then the equation is satisfied only for $Q_C > Q^*$.

Proof of Proposition 4.7. Recall from the proof of proposition 4.6 that at a symmetric equilibrium the first-order condition for each team's expected profit maximization problem can be written $Bb(n-1)/(n^2Q) + A'(Q) = w$. Thus if w is chosen so that exactly n^* teams optimally choose Q^*, then $Bb(n^*-1)/((n^*)^2Q^*) + A'(Q^*) = w$. But this implies that each team's expected profits are $E\pi = A'(Q^*)Q^* - [A(Q^*) - C] + B[n^* - b(n^*-1)]/(n^*)^2 = B[n^* - b(n^*-1)]/(n^*)^2$, which is negative if $b > n^*/(n^*-1)$. Each team has the option of choosing $Q = 0$ and earning zero profits, so choosing Q^* and earning negative expected profits is not optimal.

Proof of Proposition 4.8. In the proof of proposition 4.7, it is shown that if n^* teams compete and w is set to try to induce all teams to choose quality level Q^*, then each team would suffer expected losses equal to $B[n^* - b(n^*-1)]/(n^*)^2$. If a subsidy of $S = B[n^* - b(n^*-1)]/(n^*)^2$ is paid to each team, then each team's expected profits are zero, and there exists a symmetric equilibrium in which n^* teams each choose quality level Q^* . Similarly, if a tax of $B[n^* - b(n^*-1)]/(n^*Q_T)$ is levied on each unit of talent purchased by a team, then a total of $B[n^* - b(n^*-1)]/n^*$ in revenue is raised. If this revenue is divided equally across teams, then each team receives a tax rebate of $B[n^* - b(n^*-1)]/(n^*)^2$, expected profits are again zero, and therefore there again exists a symmetric equilibrium in which n^* teams each choose quality level Q^*.

Proof of Proposition 5.1. If the league admits n teams, then the symmetric restricted equilibrium, $Q_i = Q_T/n$ for each team $i = 1, ..., n$, and the wage is $w = Bb(n-1)/(n^2Q) + A'(Q)$ (see the proof of proposition 4.6). Substituting for Q in the wage equation yields

$w_1(n) = Bb(n-1)/(nQ_T) + A'(Q_T/n)$. Thus, the league members' joint profits are $\pi_1 = B + nA(Q) - nw_1(n)Q - nC = B + nA(Q_T/n) - w_1(n)Q_T - nC$. Pretending for the moment that n is a continuous variable, and differentiating π_1 with respect to n, the first-order condition for an interior maximum at n_1 is $A(Q_T/n_1) - C - A'(Q_T/n_1)Q_T/n_1 = w_1'(n_1)Q_T$. Also, differentiating w_1 yields $w_1'(n) = [Bb/Q_T - A''(Q_T/n)Q_T]/n^2$, which is positive, so $A(Q_T/n_1) - C - A'(Q_T/n_1)Q_T/n_1 > 0$. By definition, n^* uniquely satisfies $A(Q_T/n^*) - C = A'(Q_T/n^*)Q_T/n^*$; also, $A(Q_T/n) - C - A'(Q_T/n)Q_T/n$ is decreasing in n. Thus, $n_1 < n^*$, and since $n_1Q_1 = n^*Q^* = Q_T$, $Q_1 > Q^*$. (Note: It is possible that there is no interior solution to the league's profit maximization problem, in which case the league's optimal choice is $n_1 = 1 < n^*$. This solution is obviously unrealistic, and is possible only because of our simplifying assumptions that A and B are independent of the number of teams, and that even one team would act as a price-taker with respect to the wage.)

Proof of Proposition 5.2. Suppose each league l admits n_l teams, $l = 1, ..., L$. Then given a symmetric restricted equilibrium in each league, $Q_{li} = Q_l(\mathbf{n})$ for all $i = 1, ..., n_l$ and all $l = 1, ..., L$, $\sum_{l=1}^{L} Q_l(\mathbf{n}) = Q_T$, and the wage satisfies $w_L(\mathbf{n}) = Bb(n_l - 1)/(Ln_l^2 Q_l(\mathbf{n})) + A'(Q_l(\mathbf{n})$ for all $l = 1, ..., L$. Joint profits in each league l are $\pi_l(\mathbf{n}) = B/L + n_l A(Q_l(\mathbf{n})) - n_l w_L(\mathbf{n})Q_l(\mathbf{n}) - n_l C$, so pretending for the moment that n_l is a continuous variable, the first-order condition for an interior maximum at n_l can be written $A(Q_l) - C = -A'(Q_l)n_l\frac{\partial Q_l}{\partial n_l} + w_L Q_l + n_L Q_l \frac{\partial w_L}{\partial n_l} + n_l w_L \frac{\partial Q_l}{\partial n_l}$. If all leagues choose the same number of teams, so $n_l = n_L$ for all $l = 1, ..., L$, then $Q_l = Q_T/(Ln_L)$ for all l, $w_L(\mathbf{n}) = Bb(n_L - 1)/(n_L Q_T) + A'(Q_T/(Ln_L))$, and league l's first-order condition becomes $A(Q_T/Ln_L)) - C = [Q_T/(Ln_L)]A'(Q_T/(Ln_L)) + [Bb(n_L - 1)/(n_L Q_T)][Q_T/(Ln_L) + n_L \frac{\partial Q_l}{\partial n_l}] + (Q_T/L)\frac{\partial w_L}{\partial n_l}$. It is straightforward but tedious to show that $[Q_T/(Ln_L) + n_L \frac{\partial Q_l}{\partial n_l}]$ and $\frac{\partial w_L}{\partial n_l}$ are both positive (differentiate the first-order condition for league l, the first-order condition for any league $j \neq l$, and the market-clearing condition $\sum_{l=1}^{L} Q_l(\mathbf{n}) = Q_T$, with respect to n_l, evaluate the three equations at any $\mathbf{n}$ such that $n_l = n_L$ for all l, and solve the equations simultaneously). Thus, $A(Q_T/(Ln_L)) - C - [Q_T/(Ln_L)]A'(Q_T/(Ln_L)) > 0$. By definition, n^* uniquely satisfies $A(Q_T/n^*) - C =$

$A'(Q_T/n^*)Q_T/n^*$; also, $A(Q_T/n) - C - A'(Q_T/n)Q_T/n > 0$ is decreasing in n. Thus, $Ln_L < n^*$, and since $Ln_L Q_L = n^* Q^* = Q_T$, $Q_L > Q^*$. (Note: As in proposition 5.1, corner solutions with $n_L = 1$ cannot be ruled out. And, of course, if $L = n^*$, then $n_L = 1$ and $Ln_L = n^*$. As noted in the proof of proposition 5.1, however, this solution is quite artificial.)

Proof of Proposition 5.3. If the league admits n teams, and chooses a tax rate of τ, then the expected profit of each team i given the talent levels $Q = (Q_1, ..., Q_n)$ is $E\pi(Q) = Bp_i(Q) + A(Q_i) - (w + \tau)Q_i - C + (\tau/n)\sum_{i=1}^{n} Q_i$. At the symmetric restricted equilibrium, $Q_i = Q_T/n = Q$ for each team $i = 1, ..., n$, and the wage is $w = Bb(n-1)/(n^2 Q) + A'(Q) - \tau(1 - 1/n)$. Substituting for Q in the wage equation yields $w_1(n, \tau) = Bb(n-1)/(nQ_T) + A'(Q_T/n) - \tau(1 - 1/n)$. Thus, the league members' joint profits are $\pi_1 = B + nA(Q) - nw_1(n,\tau)Q - nC = B + nA(Q_T/n) - w_1(n,\tau)Q_T - nC$. Differentiating π_1 with respect to τ yields $\frac{\partial \pi_1}{\partial \tau} = -\frac{\partial w_1}{\partial \tau}(n,\tau)Q_T = Q_T(1-1/n)$. This is always positive, so the league chooses τ as large as possible. Given that the lowest possible wage is 0 (at wages below 0, no talent is forthcoming), this means choosing $\tau = [Bb(n-1)/(nQ_T) + A'(Q)]/(1-1/n)$, so that $w_1(n,\tau) = 0$. Given that wages are zero, total joint profits are $\pi(n) = B + nA(Q_T/n) - nC$, which is the same as total surplus (recall equation 1). Thus the league maximizes joint profits by choosing $n_\tau = n^*$ and $Q_\tau = Q^*$ (Note: This result also holds if teams take total tax revenue as fixed. The only difference is that the equilibrium wage equation is $w_1(n,\tau) = Bb(n-1)/(nQ_T) + A'(Q_T/n) - \tau$.)

Proof of Proposition 5.4. If the league admits n teams, there are two possibilities at the symmetric restricted equilibrium. If the ceiling on salaries is not binding, then the situation is exactly as in proposition 5.1, and the league's total profits are $\pi_1(n) = B + nA(Q_T/n) - w_1(n)Q_T - nC$, where $w_1(n) = Bb(n-1)/(nQ_T) + A'(Q_T/n)$. On the other hand, if the ceiling on salaries is binding, then $wQ_T/n = \theta[B/n + A(Q_T/n) - C]$, and the league's total profits are $B/n + A(Q_T/n) - wQ_T n - C = (1-\theta)[B/n + A(Q_T/n) - C]$, and the league's total profits are $\pi_\theta(n) = (1-\theta)[B + nA(Q_T/n) - nC]$. Clearly, the ceiling on salaries is binding at n if and only if $w_1(n)Q_T/n > \theta[B/n + A(Q_T/n) - C]$, that is, if and only $\pi_\theta(n) > \pi_1(n)$.

Recall from section 2 that $B + nA(Q_T/n) - nC$ is total surplus. Thus, $\pi_\theta(n)$ is equal to $(1-\theta)$ times total surplus, so $\pi_\theta(n)$ is uniquely maximized at $n = n^*$, Also, recall from proposition 5.1 that $\pi_1(n)$ is maximized at $n_1 < n^*$. Thus, there are three cases: (1) $\pi_\theta(n^*) < \pi_1(n_1)$, in which case the league's optimal choice is $n_\theta = n_1$ (the salary constraint is not binding at n_1, since $\pi_\theta(n_1) < \pi_\theta(n^*)$, and it might or might not be binding at n^*); (2) $\pi_\theta(n^*) = \pi_1(n_1)$, in which case the league is indifferent between n^* and n_1 (the salary constraint is not binding at n_1, since $\pi_\theta(n_1) < \pi_\theta(n^*)$, but it is binding at n^*, since $\pi_1(n^*) < \pi_1(n_1)$); or (3) $\pi_\theta(n^*) > \pi_1(n^*) > \pi_1(n_1)$, in which case the league's optimal choice is $n_\theta = n^*$ (the salary constraint might or might not be binding at n_1, but it is binding at n^*, since $\pi_1(n^*) < \pi_1(n_1)$). Thus, in order to guarantee that the league chooses $n_\theta = n^*$, θ should be set so that $\pi_\theta(n^*) > \pi_1(n_1)$, or, letting S^* be the maximized value of total surplus, $\theta < 1 - \pi_1(n_1)/S^*$.

Proof of Proposition 6.1. Total Social surplus is $B(b) + nA(Q_T/n) - nC$, so it is easily seen that this is maximized by choosing $(b, n) = (b^*, n^*)$, where b^* solves $B'(b^*) = 0$ and n^* satisfies $A(Q_T/n^*) - C = A'(Q_T/n^*)Q_T/n^*$.

Suppose the league chooses degree of determinism b, and admits n teams. Then at the symmetric restricted equilibrium, $Q_i = Q_T/n$ for each team $i = 1, ..., n$, the wage is $w_1(n) = B(b)b(n-1)/(nQ_T) + A'(Q_T/n)$, and the league members' joint profits are $\pi_1 = B(b) + nA(Q_T/n) - w_1(n)Q_T - nC$. Differentiating π_1 with respect to b and n (again, we ignore for the moment that n is discrete), the first-order conditions for an interior maximum at (b, n_b) are $B'(b) = B(b)(n_b-1)/[n_b - b(n_b-1)]$, and $A(Q_T/n_1) - C - A'(Q_T/n_1)Q_T/n_1 = [B(b)b - A''(Q_T/n)Q_T^2]/n^2 > 0$. Also, differentiating π_1 twice with respect to b and evaluating the result at (b_b, n_b) yields

$$\frac{\partial \pi_1^2}{\partial b}(b_b, n_b) = \frac{B''(b_b)[n_b - b_b(n_b-1)]^2 - 2(n_b-1)B(b_b)}{n_b[n_b - b_b(n_b-1)]}$$

The numerator of this expression is clearly negative, so $\frac{\partial \pi_1^2}{\partial b}(b_b, n_b) < 0$ if and only if the demoninator is positive, which holds if and only if $b_b < n_b/(n_b - 1)$. Substituting this back into the first-order conditions yields $B'(b_b) > 0$, which implies that $b_b < b^*$. Finally, the proof that $n_b < n^*$ follows exactly the same logic as the proof of proposition 5.1 (use the fact that $A(Q_T/n) - C - A'(Q_T/n)Q_T/n$ is decreasing in n).

References

Canes, Michael E. 1974. "The Social Benefits of Restrictions on Team Quality." In *Government and the Sports Business*, edited by Roger Noll. Washington, D.C.: Brookings Institution.

El Hodiri, Mohamed, and James Quirk. 1971. "An Economic Model of a Professional Sports League." *Journal of Political Economy* 79: 1302-19.

Friedman, L. 1958. "Game-Theory Models in the Allocation of Advertising Expenditures." *Operations Research* 6: 699-709.

Quirk, James, and Mohamed El Hodiri. 1974. "The Economic Theory of a Professional Sports League." In Roger Noll, ed., *Government and the Sports Business*. Washington, D.C.: Brookings Institution.

McCormick, Robert E. and Robert Clement. 1991. "Infrafirm Profit Opportunities: Evidence from Profession Basketball." In Gerald W. Scully, ed., *Advances in the Economics of Sports*. New York: JAI Press.

Noll, Roger. 1974. "Attendance and Price Setting." In Roger Noll, ed., *Government and the Sports Business*. Washington, D.C.: Brookings Institution.

Rosen, Sherwin. 1986. "Prizes and Incentives in Elimination Tournaments." *American Economic Review* 76: 701-709.

Scully, Gerald W. 1989. *The Business of Major League Baseball*. Chicago: University of Chicago Press.

Snyder, James M., Jr. 1989. "Election Goals and the Allocation of Campaign Resources." *Econometrica 57*: 637-660.

Tullock, Gordon. 1980. "Efficient Rent-Seeking." In *Towards a Theory of the Rent-Seeking Society*, edited by J. Buchanan, R. Tollison, and G. Tullock. College Station: Texas A&M University Press.

Whitney, James D. 1993. "Bidding Till Bankrupt: Destructive Competition in Professional Team Sports." *Economic Inquiry* 31: 100-115.

Figure 1

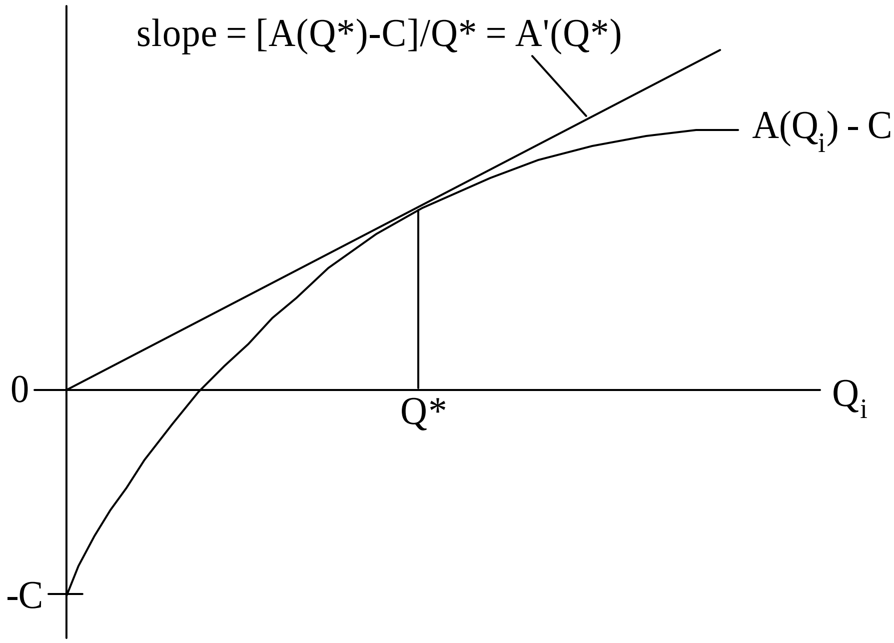

Figure 2