A General Theory of Professional Sports Leagues

JOHN VROOMAN
Vanderbilt University
Nashville, Tennessee

The evolution of free agency in professional sports leagues spans two decades from the 1972 Flood v. Kuhn decision in Major League Baseball (MLB) to the 1992 McNeil, et al. v. NFL decision in the National Football League (NFL). In spite of the apparent relaxation of eligibility requirements for free agency, ancillary institutional arrangements have severely limited player mobility during these two decades in all leagues but MLB. As part of a “revolutionary partnership,” National Basketball Association (NBA) players agreed to a league-wide payroll cap for team salaries in 1984 in exchange for a guaranteed 53 percent of NBA designated gross revenues (DGR). The ostensible purpose of the NBA payroll or “salary” cap was to control spiraling salaries and generate competitive balance within the league, but since its inception, the payroll cap has served to limit player mobility under free agency, and its effect on competitive balance is subject to question. Although free agency eligibility requirements have been lenient in the NFL, strict compensation rules have thwarted player movement until McNeil. In spite of free agency concessions by NFL owners following McNeil, movement of players among teams will continue to be limited under The Collective Bargaining Agreement of 1993, because of the concurrent imposition of a payroll cap at 64 percent of league-wide DGR beginning with the 1994 season.

1. Flood v. Kuhn, 107, U.S. 258 (1972). The decision did not address the right of a MLB Club to control the rights of a player (Curt Flood of the St. Louis Cardinals) for perpetuity. It concerned the continuation of an antitrust exemption enjoyed by MLB in two previous decisions: Toolson v. New York Yankees, 346 U.S. 356 (1953); and Federal Baseball Club v. The National League, 259 U.S. 200 (1922). In casting the deciding vote, Chief Justice Burger admitted that the Court was originally in error, but that it was such an old error that it was up to Congress to correct the mistake. The question of free agency would remain unanswered until the Collective Bargaining Agreement of 1976, in which the limits of the reserve clause were reduced to six years.


3. This paper has been significantly improved by the candid suggestions of an anonymous referee. One such recommendation concerns calling the “salary cap” a payroll cap, because the so-called “salary caps” actually limit team payroll not individual player salaries. The candor of this referee is appreciated.

4. The Supreme Court made it clear that only MLB enjoyed the antitrust exemption in Radovich v. NFL, 352 U.S. 445 (1957). NFL players were eligible for free agency after the expiration of the option year in their contracts but, according to the “Rozelle Rule,” a team signing a free agent was required to compensate the original team by means of players or draft selections. The Rozelle rule was set aside as an unreasonable restraint of trade in Mackey v. NFL, 543 F. 2nd 644 (8th Cir. 1976), Cert. dismissed, 434 U.S. 801 (1977). The NFL’s “Plan B” free agency system, where a team could protect 37 of 47 roster players, was rendered illegal in McNeil, et al., v. NFL, 790 F. Supp. 871 (8th Cir. 1992). Following McNeil NFL management and the NFLPA reached an accord in 1993 which granted unrestricted free agency in the NFL.
During the second decade of the free agency era, the payroll cap has coevolved with free agency as its countervailing companion, and the payroll cap has simultaneously emerged as the major, unresolved issue in the collective bargaining of all professional sports leagues.5

The arguments for the payroll cap and strict compensation rules are the same as those traditionally made for the reserve clause. First, it is argued that player salaries should be controlled so that teams could recoup their investment in the development of talent. Second, it is maintained that talent will gravitate toward large market teams under unbridled free agency, and that constraints are necessary to defy gravity and maintain competitive balance within the league. In opposition to these arguments, conventional economic theory has held that the distribution of talent would be the same under free agency as it was under the reserve clause, and that competitive balance would not be affected by institutional change. This argument was originally made by Rottenberg when he surmised “that a market in which freedom is limited by a reserve rule such as that which now (1956) governs the baseball labor market distributes players among teams about as a free market would [25, 255],” and later embellished by Demsetz as application of the “Coase theorem” [4] when he asserted: “No matter who owns the right to sell the contract for the services of a baseball player, the distribution of players among teams will remain the same [10, 17].” The invariance proposition was formalized in the work of Quirk and El Hodiri [11; 23], and recent studies have reached a consensus that there has been no change in competitive balance due to free agency in either MLB or the NBA.6

Thus, while the legal and institutional foundations of major sports leagues have been coevolving in the era of free agency, conventional economic theory has not. Daly observes that “Rottenberg’s invariance proposition proved compelling to so many economists, some of whom viewed its logic to be so unassailable as to constitute a proof of its validity. Its grip on economists thinking persists to this day [6, 14].” The purpose of this paper is to reconsider the implications of conventional economic theory within the institutional configurations of professional sports leagues as they have evolved. This immediately involves the reduction of conventional theory to a special case of a more general theory of sports leagues.

I. A General Theory

The economics of professional sports has been preoccupied with the dual proposition that a large market team will dominate a small market team, and that the competitive imbalance will be invariant under a variety of institutional constraints designed to alter it. Quirk and El Hodiri [11; 23] developed an elegant model of a sports league in which: “The evidence points to the conclusion that the rules structure of professional sports is relatively ineffective in balancing playing strengths, and that the imbalance is due to the differences in the drawing potentials of franchises [23, 58],” or more simply “big cities have winning teams and small cities have losing teams [23, 45].” The purpose of this section is to reconsider the implications of conventional theory for

5. Unresolved disputes concerning the payroll cap have resulted in the cancellation of the 1994 MLB season and have put the 1994 National Hockey League season “on ice.” Attempts to harden the “soft” (avoidable) NBA cap by owners and antitrust litigation against the cap by players have fragmented the “revolutionary partnership” and threatened the 1994 season. NFL players are confused about the chilling effect of “cap-onomics” on their long-awaited free agency; and the unilateral implementation of the salary cap by MLB owners has again stirred Congressional reconsideration of the unique antitrust exemption continued in Flood.

6. See Scully [26], Balfour and Porter [2], and QF [24]. Some early studies argued that the reserve clause increased equality of competition in MLB, see Daly and Moore [7], Cymrot [5] and Lehn [15].
the contemporary world of professional sports, and to explore the possibility that large market dominance might be negated by other economic and institutional characteristics of professional sports leagues and the teams that comprise them.

The implications of conventional economic theory can be seen through discussion of a model of a simplified professional sports league with two teams, each with the revenue and cost functions:

\[ R_i = R_0 p_i \gamma w_i^{\beta} \]
\[ C_i = C_0 p_i \gamma w_i^{\delta} \]

where:
- \( R_i \) is the total revenue for team \( i \)
- \( R_0 \) is exogenous revenue for both teams
- \( \alpha \) is the market size revenue effect
- \( \beta \) is the revenue elasticity of winning
- \( p_i \) is the home market size for team \( i \)
- \( C_i \) is the total cost for team \( i \)
- \( C_0 \) is exogenous cost for both teams
- \( \gamma \) is the market size cost externality
- \( \delta \) is the cost elasticity of winning
- \( w_i \) is the winning percentage for team \( i \)

and the zero-sum restrictions of the league require:

\[ \sum_{i=1}^{n} w_i = n/2 = 1.000. \]

The profit maximization conditions for each team yield the general competitive balance equilibrium:

\[ w_1/w_2 = \left( p_1/p_2 \right)^{(\alpha-\gamma)/(\delta-\beta)}. \] (1)

The customary assumptions of conventional theory can be summarized for the parameters in this model:

**Alpha**: Market size creates a revenue advantage that results in large market dominance on the playing field, i.e., \( \alpha > 0 \) for \( p_1 > p_2 \).

**Beta**: Marginal revenue is assumed to be a nonincreasing function of winning and all teams are assumed to have similar revenue functions, i.e., \( \beta \leq 1 \).

**Gamma**: The price per unit of talent is assumed to be the same for all teams, i.e., \( \gamma = 0 \).

**Delta**: Marginal cost is assumed to be a nondecreasing function of winning and all teams are assumed to have similar cost functions, i.e., \( \delta \geq 1 \).

The model bridges two separately incomplete treatments of competitive balance found in the recent literature: Quirk and Fort [24] and Scully [26]. Quirk and Fort (QF) base their rendition of Quirk's earlier work with El Hodiri (QE) on the assumption that "both teams will face the same market cost per unit of playing strength, and hence the same cost to increase the team's win/loss percentage (emphasis added) [24, 273]." This assumption can only be valid if there are no externalities of market size and if marginal costs are constant and the same for each team regardless of its winning percentage. QF's model amounts to a special case of (1):

7. The strength of the two-team model derives from its simplicity and efficiency in dealing with the questions of talent polarization. An anonymous referee observes that the two-team model may be unable to deal with the dominance diseconomies for third and fourth teams in the league. The importance of winning among nondominant teams may be diminished unless they are in contention with the dominant team. This would imply that the revenue elasticity of winning \( \beta \) is an inverse function of dominance, and that the competitive dominance implied by (1) is understated. Although a more complex four-team model could capture the externality, *Occam's razor* favors the simple two-team model.

8. QE make the assumption as well: "All franchises are assumed to pay the same wage cost per unit and the same prices in the existing market and the draft... Under these assumptions, the marginal cost of acquiring an additional unit of playing skills is the same for every team, whatever its playing strength [23, 38]."
Figure 1. The Invariance Proposition

$$w_1/w_2 = [p_1/p_2]^{\alpha/(1-\beta)} \quad (2)$$

where $\gamma = 0$ and $\delta = 1$ by QF assumption. This competitive balance equilibrium is shown (as presented in QF) in Figure 1 for the large market Team 1 (left to right) and the small market Team 2 (right to left). Under the constant marginal cost assumption of QF, playing talent will move from Team 2 to Team 1 for any competitive balance less than .600/.400. Equal playing strengths between Team 1 at B and Team 2 at C create a disequilibrium condition because the potential gains to Team 1 outweigh the losses to Team 2 by the triangle ABC. All potential gains from talent redistribution for either team will be exhausted at A. The alpha factor revenue advantage enjoyed by Team 1 results in the league competitive balance equilibrium at A, and its .600/.400 dominance of Team 2.

Following the invariance proposition, QF observe that A in Figure 1 will be the league equilibrium condition, regardless of whether the league is operating under free agency or such institutional player restrictions as the reserve clause or a salary cap. According to the invariance proposition, the economically justifiable distribution of talent in a Coasian world of zero transactions costs is independent of the legal ownership of that talent. The difference, of course, is that under free agency the players' salaries are commensurate with A, whereas under the reserve clause (or a salary cap), Team 2 will capture the gains and players will be paid less than the value of their marginal product at B. Conventional theory implies that the large market team will dominate the small market team with or without the institutional constraints on free agency, and that the difference under either regime relates not to the distribution of talent, but to the distribution of income derived from that talent.$^9$

Clearly the implications of (2) are that competitive balance is determined, not only by $p_1 > p_2$ and $\alpha > 0$, but also by the magnitude of $\beta$, the revenue elasticity of winning. Solution (2) suggests the following beta factor elasticity propositions and attendant corollaries:

9. A small market “sportsman owner” may include his preference for winning in his objective function and “consume” talent and compete with the large market Team 1 at C in Figure 2. In this case marginal costs exceed marginal revenue and the team’s long term existence depends ultimately on the wealth of the owner.
Figure 2. The Fickle Fan Effect

**Proposition 1.** If the revenue elasticities of winning are equal for the two teams, then competitive balance in the league will decrease and the large market team will become increasingly dominant as the revenue elasticity of winning increases. This is the move from A to B in Figure 2.

**Corollary.** If the revenue elasticities of winning for each team are equal but decreasing, competitive balance will increase but, as long as market size is a significant factor, the league will not reach complete parity. This is shown as a move from B to A in Figure 2.

**Corollary.** An increase (decrease) in either the large market advantage \( \alpha \) or the revenue elasticity of winning \( \beta \) for either or both teams increases (decreases) marginal revenue at equilibrium and if players receive the value of their marginal product, this increases (decreases) player costs. This is shown at equilibria A and B in Figure 2.

Thus it seems that, while the level of dominance or competitive imbalance in sports leagues may vary depending on the revenue elasticity of winning, the large market team will always dominate the small market team to some degree. This conclusion depends, however, on the assumption of equal revenue elasticities of winning for the teams in the league. The solution of (2) for \( \beta_1 \neq \beta_2 \) implies:

**Proposition 2.** Even if market size is a significant determinant of team revenue, a small market team can still be competitive if its fans have a sufficiently higher elasticity of the demand for winning than do the fans in the large market. The competitive balance of the relatively elastic small market team is shown at C and the increased dominance of the relatively elastic larger market team at D in Figure 2.

Another important factor that has been ignored is the effect of increasing marginal costs of playing talent on competitive balance. Hunt and Lewis (HL) make the reasonable point that

10. This fickle-fan proposition is made independently by Porter: “Seen in this light fans have a larger stake in determining the quality of their team, not by being loyal, but by being fickle. The more elastic the attendance response to wins, the greater the incentive of the owner to field a winning team” [22, 64].
“the amount that the dominating team must offer for the player resources necessary to increase dominance is equal to the loss in revenue that the other teams would suffer due to the resource transfer [13, 940].” In this case the marginal cost curve for Team 1 is seen as the marginal revenue curve for Team 2 in Figure 1. Team 1 would continue to acquire talent until all potential gains are exhausted at profit maximization equilibrium A in Figure 1. Although this is a slightly different argument than that of QF, it yields the same competitive equilibrium. HL’s position is plausible, but like the argument of QF, it is a special case and it unnecessarily limits the inquiry into the effects of free agency on competitive balance.

Implicit in HL and explicit in QF are the joint assumptions of a constant marginal cost of talent and a constant marginal productivity of talent. Diminishing marginal returns to talent would imply that marginal costs for the two teams are symmetrical and increasing, and that marginal costs would be equal for the two teams only when the winning percentages of the two teams are equal at .500/.500. If the marginal product of playing talent with respect to winning is diminishing, then the marginal product of playing talent acquired by Team 1 from Team 2 will be less for Team 1 than it was for Team 2 for the pursuit of winning percentages above .500. Therefore, unless the dominant team faces a significantly lower wage per unit of talent through economies of market size, the marginal costs for the dominant team will be higher than the marginal revenue for the dominated team for any winning percentage above .500. As a result, the actual competitive solution under profit maximization will be more balanced than that predicted by QF’s league revenue maximization solution at A in Figure 1. Increasing marginal costs in the competitive balance solution (1) imply \( \lim_{k \to \infty} w_1 = w_2 = .500 \) and yield the \( \text{delta} \) factor proposition:

**Proposition 3.** Competition becomes more (less) balanced as the cost elasticity of winning increases (decreases).

Furthermore, it is not necessary to limit the argument with the assumption that both teams are subject to the same wage rate per unit of talent. Demmert notes that “the empirical question of what particular market variations affect a club’s costs of acquiring athletic talent is of some interest [9, 51].” He suspects that the commercial opportunities in the larger markets might supplement players’ incomes and reduce both the wage rate and the marginal costs for those teams in large markets. Although endorsement opportunities create positive externalities of market size (\( \gamma < 0 \)), recent difficulties in the free agent signings of large market MLB franchises may suggest renewed importance of the diseconomies of a market size (\( \gamma > 0 \)) in the allocation of playing talent.\(^{11}\) If the large market team has diseconomies of market size and both teams face nondecreasing marginal costs, then it will have higher marginal costs at .500, and competition with the small-market team will be more balanced that suggested by QF’s analysis.\(^{12}\) The general competitive balance solution derived above in (1) implies \( \lim_{\gamma \to -\infty} w_1 = w_2 = .500 \) and yields the \( \text{gamma} \) factor proposition:

**Proposition 4.** Competition between two teams will become more balanced as diseconomies of market size increase and compensate for revenue advantages derived from market size.

\(^{11}\) New York Yankee General Manager Gene Michael estimates that free agents seek a ten to fifteen percent surcharge on Yankee contract offers. Following the 1992 season, the Yankees pursued six top free agents and signed none. Two of the six, Greg Maddux and Doug Drabek, signed contracts with other teams for almost $1 million per year less than the Yankees’ best offer. See Nightingale [17].

\(^{12}\) QE footnote Walter Ol’s suggestion that, “to the extent that wage rates are higher in larger cities, the difference in playing strengths between large- and small-city teams is reduced; and if the difference in wage rates were significantly great, one might even find that the small cities used larger stocks of playing skills than large cities [23, 66].”
Figure 3. Free Agency and the Marginal Cost of Talent

The competitive balance consistent with revenue maximization for the league will yield a profit maximum for each team, if and only if, marginal costs are constant and there are no cost externalities due to market size, i.e., $\delta = 1$ and $\gamma = 0$. The more general case where both teams face the same diminishing marginal product of talent ($\delta > 1$), and Team 1 faces diseconomies of market size ($\gamma > 0$) is shown in Figure 3. Compare the competitive balance solution under revenue maximization at $A$ with the more general solution under profit maximization for Team 1 at $B$ and Team 2 at $C$. Clearly the omission of the increasing marginal cost of talent biases the perception of the competitive dominance of large market franchises.

II. The Revenue Sharing Paradox

In contrast to QF, Scully [26, 78–9] allows for the increasing marginal cost of talent, but overlooks the zero-sum restrictions of the league. As a result, the unconstrained Scully model does not have a unique competitive solution, and while it does yield qualitatively accurate descriptions of large market dominance, the omission leads to the oversight of the effects of the negative interdependence of winning and losing. Consequently, the model misleads Scully to the widely accepted conclusion that “a change to an even share in the gate split would redistribute revenues from the big city team to the small city team. Under a 50-50 gate split the marginal revenues of the two teams would be identical. Hence, the win percentages would tend toward equality [26, 80].” Although this proposition is intuitively appealing, it is not necessarily true.

Attempts to improve competitive balance through various forms of revenue sharing may or may not have the intended result depending on the winning elasticity of the revenue source that is shared. A major consideration that is often overlooked in detecting the impact of elastic

---

13. The gamma factor externality of market size is the difference between the marginal costs for the two teams at .500. Equilibria $B$ and $C$ will still be more balanced than $A$ even in the absence of the externality ($\gamma = 0$), and the relative balance will be further diminished by a positive externality of market size ($\gamma < 0$). Joint profit maximization for two teams under collusion would depress player costs but, due to the negative interdependence of winning and losing, collusion would have no impact on competitive balance.
revenue sharing is the negative interdependence between the teams’ marginal revenue (and cost) functions implied by the zero-sum existence of the league: \( \partial w_1 / \partial w_2 = \partial w_2 / \partial w_1 = -1 \). Under winning-elastic revenue sharing where \( \lambda \) is the home team’s revenue share, the internalization of the interdependence alters the marginal revenue functions of both teams:

\[
MR'_1 = \lambda MR_1 - (1 - \lambda)MR_2 \\
MR'_2 = \lambda MR_2 - (1 - \lambda)MR_1
\]

where each function contains a term allowing for the negative interdependence. Paradoxically, revenue maximization with revenue sharing at \( MR'_1 = MR'_2 \) yields the same competitive balance solution as (1) without revenue sharing \( MR_1 = MR_2 \). The negative interdependence creates this paradoxical result because, when a team shares revenue, the visitor’s share depends directly on the demand for winning of its opponent’s fans, which varies inversely with its own ability to win on the road. Under the conditions of winning-elastic revenue sharing, it pays a team to win at home and lose (so that its opponent can win) on the road. Thus, while mutual winning-elastic gate sharing does not necessarily make the marginal revenue functions of the teams identical, as Scully argues, it does affect the winning incentives of the teams equally. This leads to the conclusion that if all teams are subject to the same sharing formula, then the sharing of winning-elastic revenue has no effect on competitive balance.\(^\text{14}\)

The important differences between the effects of revenue maximization with or without elastic revenue sharing lies, not in the competitive balance solution, but in the value of marginal product and player costs at equilibrium. If \( MR'_1 = MR'_2 \) implies the same competitive balance as \( MR_1 = MR_2 \), then winning-elastic revenue sharing under \( MR'_1 = \lambda MR_1 - (1 - \lambda)MR_2 \) implies \( MR'_1 = MR'_2 = (2\lambda - 1)MR_1 \). If players are paid the value of their marginal product then, for any \( 0.5 < \lambda \leq 1 \), winning-elastic revenue sharing implies the exploitation of players, \( MC_i = MR'_i < MR_i \). The mutual sharing of winning-elastic revenues results in lower marginal revenues for all teams and lower player costs at equilibrium, and as the home teams share \( \lambda \) becomes smaller, the exploitation of players increases. The effects of winning-elastic mutual revenue sharing for decreasing home team shares are shown at A, B and C in Figure 4.

It seems reasonable to argue that the one-way transfer of large market media revenues to small markets would increase competitive balance. This is not necessarily true, because of the negative interdependence of team revenues. For example, Noll concludes that the collusive sharing of national media revenues results in big city dominance in the NBA. “In reality, parity is unlikely, for it is not the financial interest of the league. A good team will do better financially and contribute more to the value of a league’s national broadcast package if it is in a big city . . . teams in smaller markets will derive financial benefits from the success of the teams in big markets by sharing in fatter national broadcast packages [20, 33–4].” Thus, while one-way winning-elastic revenue sharing intuitively implies greater competitive balance, Noll implies that it would foster less balance.\(^\text{15}\)

The question can be resolved with the realization that, while Noll’s intuitive consideration of the negative interdependence is correct, it is nonetheless incomplete. Under a winning-elastic revenue sharing formula where \( \lambda \) is the large market team’s share, the marginal revenue functions for the teams become:

\(^\text{14}\) This argument was developed in QE [23] and is well articulated in QF [24].

\(^\text{15}\) Noll’s argument implicitly assumes that media revenue is winning elastic, otherwise it would not matter whether or not the winning team was in the large market. Zimbalist [30] makes the same argument for MLB.
Figure 4. The Revenue Sharing Paradox

\[ MR'_1 = \lambda MR_1 = MR_1 - (1 - \lambda)MR_1 \quad MR'_2 = MR_2 - (1 - \lambda)MR_1 \]

where \( MR'_1 = MR'_2 \) implies the same competitive balance as the solution without revenue sharing: \( MR_1 = MR_2 \). The marginal revenue function of the small market Team 2 allows for the negative interdependence of the small market team's revenue with the one-way revenue share that depends, in turn, on the success of the large market team. Noll and others overlook the identical disincentive effect that the one-way transfer has on the marginal revenue of the large market Team 1. This leads to a statement of the revenue-sharing paradox that holds true for either mutual or one-way arrangements:

**Proposition 5.** The sharing of winning-elastic revenues will have equal disincentive effects on all teams and, while it will effectively depress player costs, it will not affect competitive balance. One-way revenue sharing from large market to small market teams will increase competitive balance, if and only if, the revenue that is transferred is winning-inelastic.\(^{16}\)

III. The Myth of the Salary Cap

Another attempt to improve competitive balance in sports leagues involves the imposition of a league-wide maximum cap on team payrolls. In their analysis of the effects of a “salary cap” on competitive balance, QF surmise that “if all teams end up spending an equal amount equal to the salary cap, then . . . the league would end up with all teams having roughly the same playing

---

\(^{16}\) A nontrivial ambiguity arises with the possibility that an increase (decrease) in the share of revenue which is winning inelastic may decrease (increase) the elasticity of winning. In this case the revenue advantage gained through alpha factor sharing may be negated through beta factor feedback. For example, MLB owners reached an agreement in 1994 (contingent on the payroll cap) in which the top third highest revenue teams agreed to share revenue with the bottom third (the middle revenue tier would neither give nor receive). While this may compensate for the large winning-inelastic revenue advantage, it could increase the financial dependence of such teams as the New York Yankees and Mets ($47 million and $37 million in local media revenues compared to $5 million to $10 million for the small revenue teams) on the actual on-field performance of their teams.
strength" [24, 287]. According to QF, the cap would have an effect on salaries similar to the reserve clause as shown in Figure 1, but because of player immobility under the cap, the league balance would remain at C. Unfortunately, QF’s view of the salary cap assumes something to be true which has yet to be proven. The purpose of the cap may not be to achieve competitive balance. Instead, the cap could serve as a collusive attempt to control total player costs, and it would allow the maximization of profits for the league as a whole. Under a payroll or “salary” cap, all teams share equally the player costs that increase with league revenues from year to year, regardless of the winning percentages of the individual teams. Consequently the marginal costs of winning for the individual teams are equal to zero, and profits for the individual teams will be maximized at the league revenue maximum shown at A in Figures 1 and 3. Second, it does not necessarily follow that equal spending among teams implies equal playing strengths among those teams. Under a salary cap and constant marginal costs the distribution of talent within the league may not be affected by the diminishing marginal productivity of talent (δ = 1), but it is still a function of the revenue characteristics (α and β) of the teams as well as the (dis)economies of the home markets (γ ≠ 0) in which the teams play. Equation (1) can be rewritten for league profit (revenue) maximization under the salary cap where δ = 1:

\[ \frac{w_1}{w_2} = \left[ \frac{p_1}{p_2} \right]^{(\alpha-\gamma)/(1-\beta)} \]  

(3)

The imposition of a payroll cap effectively eliminates increasing marginal costs of talent (δ = 1) from the profit maximization calculus, and it thereby allows the teams of the league to collusively behave as the firm. As a result, the capping of team payrolls ironically promotes competitive imbalance within the league, rather than the competitive balance surmised by QF. Compare the collusive payroll cap solution A with the free agency competitive solutions B and C in Figure 3.

IV. The Free Agency Illusion

The application of this general theory to the “reality” of free agency requires a discussion of the institutional configurations of the three leagues. A summary of these constraints is presented in Table 1. Since the Basic Agreement of 1976, the eligibility requirement for free agency has remained the same in MLB: “Any player with 6 or more years in Major League Service who

17. QF observe that the cap will create enforcement problems as the teams seek revenue maximization at A.
18. Given constant player costs, talent will still migrate within the league to those markets which promised the highest non-playing income possibilities. Michael Jordan’s salary of $3.9 million from the NBA Chicago Bulls in 1992–93 was supplemented by approximately $32 million in endorsements.
19. An interesting variation of the payroll cap that was suggested in both the MLB and NHL negotiations of 1994 involves a luxury tax or surtax (two to five percent) on the amount of a large revenue team’s payroll that exceeds a predetermined limit (cap). The surtax is pooled and redistributed to the small revenue teams. Because of negative interdependence, the incidence of the tax would adversely affect the winning incentives of the teams equally. Small revenue teams would be discouraged from winning so as to share in the subsidy, while large market teams would be discouraged equally by the tax itself. The surtax, however, would still result in greater competitive balance than the conventional payroll cap, but less balance than under free agency. This is because large revenue teams would encounter increasing marginal costs, whereas small revenue teams would face zero marginal costs. The surtax competitive equilibrium of \( MR_1 - MC_1 = MR_2 \) would lie between A and B/C in Figure 3. Due to disincentive effects, the tax would also allow exploitation of players. Small revenue teams would pay under the cap, whereas player costs for large revenue teams would be \( MC_1 = MR_1/(1+t) \) for the tax rate t.
| **Table I. Current Institutional Arrangements of Professional Sports Leagues** |
|---------------------------------|---------------------------------|---------------------------------|
| **Bargaining Agreement:**  | **MLB\textsuperscript{a}** | **NBA\textsuperscript{b}** | **NFL\textsuperscript{c}** |

<table>
<thead>
<tr>
<th><strong>Free Agency:</strong></th>
<th><strong>MLB\textsuperscript{a}</strong></th>
<th><strong>NBA\textsuperscript{b}</strong></th>
<th><strong>NFL\textsuperscript{c}</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Moderate: unrestricted after 6 years with draft picks as compensation.</td>
<td>Liberal: unrestricted after expiration of 1st contract with right of first refusal to original team. No right of first refusal after expiration of 2nd contract after 5 years (after 4 years in 1993).</td>
<td>Liberal after 1993: unrestricted after 4 years except one franchise player. Strict before 1993: unconditional for Plan B unprotected players, conditional after 4 years with draft compensation.</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Payroll Cap:</strong></th>
<th><strong>MLB\textsuperscript{a}</strong></th>
<th><strong>NBA\textsuperscript{b}</strong></th>
<th><strong>NFL\textsuperscript{c}</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>None</td>
<td>Soft: 53 percent of league national media and gate revenues. Media is exact and 10 percent added to gate of prior year; may be exceeded to resign own free agent.</td>
<td>Hard after 1993: 64 percent of league gate and national media; rookie cap is based on team's position in the draft. None before 1994.</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Gate Sharing:</strong></th>
<th><strong>MLB\textsuperscript{a}</strong></th>
<th><strong>NBA\textsuperscript{b}</strong></th>
<th><strong>NFL\textsuperscript{c}</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>(Home/Visitor) Gate is 33 percent of total revenue and is split 80/20 in AL and 90/10 in NL.</td>
<td>Gate is 42 percent of total revenue and is not shared (100/0).</td>
<td>Gate is 24 percent of total revenue and is split 60/40.</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Media Sharing:</strong></th>
<th><strong>MLB\textsuperscript{a}</strong></th>
<th><strong>NBA\textsuperscript{b}</strong></th>
<th><strong>NFL\textsuperscript{c}</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Media is 44 percent of total revenue and 60 percent of media revenue is shared national; 20 percent of AL local media is shared and 25 percent is shared in NL.</td>
<td>Media is 40 percent of total revenue and 67 percent of total media is shared national. Local media is not shared.</td>
<td>Media is 68 percent of total revenue and 95 percent of total media is shared national. Local media is not shared.</td>
<td></td>
</tr>
</tbody>
</table>

\textsuperscript{a} Under the MLB media contract for 1994–99 with NBC, ABC, and ESPN, a team's annual share of national revenue will drop from $14.5 million per year in 1990–93 to an estimated $6.4 million for each of the 28 teams (including 2 expansion teams).  
\textsuperscript{b} NBA per team national media revenues should rise from $8.1 million annually in 1990–93 (NBC, TNT) to $10.2 million in 1994–97 (NBC, TNT, TBS).  
\textsuperscript{c} NFL per team national media revenues are expected to increase from $32.5 million in 1990–93 (CBS, NBC, ABC, ESPN, TNT) to approximately $38 million in 1994–97 (FOX, NBC, ABC, ESPN, TNT).

has not executed a contract for the next succeeding season shall be eligible to become a free agent [3, XX(B)(1)]. Major changes that may have affected free agent mobility since then have involved the type of compensation (draft picks or players) due to the team that loses the free agent. The eligibility requirement for the NBA has also remained the same since its merger with the American Basketball Association in 1976. A NBA player is eligible to become an unrestricted free agent after the expiration of his first contract. Major free agency changes in the NBA since the merger have involved relaxation of the compensation rule (eliminated 1980) and changes in
the right of first refusal for the original team. From 1977 to 1993, the NFL operated under the illusion of a lenient free agency eligibility requirement that a player who has completed his second year in the NFL and played out the option year of his contract would have become a “conditional free agent.” In spite of the lenient eligibility requirement, free agency for NFL players during this period was nominal because of the extremely severe conditions of compensation to the original team. After the Collective Bargaining Agreement of 1993 a player can become an unrestricted free agent after four years service in the NFL.

The ultimate impact of free agency on competitive balance and player costs has been tempered, not only by restrictive compensation clauses, but also by the joint implementation of payroll caps. Both of the agreements by players’ unions to cap salaries have involved compromises by owners to fix the cap at a constant portion of league-wide designated gross revenues (usually gate and national media revenues). NBA team owners and players agreed to cap salaries starting in 1984 at 53 percent of league-wide DGR. In a similar compromise, the NFLPA agreed to cap salaries starting in 1994 at 64 percent of DGR in exchange for free agency after four years. MLB has no payroll cap at present, and the unresolved dispute between MLB owners and the MLBPA over the implementation of a salary cap resulted in the players’ strike and the cancellation of the 1994 season.

Although the magnitude of the revenue elasticity of winning appears to be an empirical question, there is sufficient variation among the institutional configurations of the leagues to allow the formulation of hypotheses a priori. Based on the work of Horowitz [12], one would suspect that national media revenues would be winning-inelastic with respect to an individual team. It follows that the winning-elasticity of revenue would be inversely related to the share of total team revenue derived from shared national media. In 1992, approximately two thirds of media revenues were shared in both MLB and the NBA, while 95 percent of media revenues were shared in the NFL. Media revenues are less than one half of total revenues in MLB and the NBA, while media revenues are two thirds of the total revenue of an NFL team. Therefore, approximately one third of the total revenue in MLB and the NBA is inelastic shared media revenue, whereas approximately two thirds of total NFL revenue is derived from winning-inelastic shared media sources. Based on this media-sharing hypothesis, the NFL should have the most winning-inelastic revenues and lowest player costs of the three leagues.

Another league-specific factor that might affect the revenue elasticity of winnings is the length of the season. If each home game is seen as a potential substitute for another, then the winning elasticity of attendance and gate revenues would be a direct function of the number of home games. In a typical regular season, each MLB plays 81 home games, each NBA team plays 41 home games (indoors), and each NFL team plays eight home games. The predictions of

20. Since 1988 right of first refusal remains with the team after expiration of the first contract. There is no such right after expiration of the second contract for 5 year veterans (4 year veterans in 1993–94).

21. Depending on the salary and years experience of the player, the original team retained first refusal rights and could receive up to two first round draft picks as compensation for the lost player. During the twenty-six years of nominal free agency, only two players changed teams via these conditions.

22. The NFL Management Council [16] unilaterally imposed “Plan B” modifications to free agency after 1987, which were struck down in McNeil in 1992. In the 1993 accord, each team could protect a designated “franchise” player; in the transition year 1993 the eligibility requirement was five years.

23. The NBA cap may be exceeded for a variety of “grandfather” exceptions. For example, a team may exceed the “soft” cap to resign its own free agents, and to extend the contracts of its veterans. A team that is over the cap can continue to exceed the cap to replace players who have left via retirement, injury or free agency by offering the replacement player 50 percent of the departed player’s salary. See D’Alessandro [8]. In contrast, the NFL cap is hard in that it can not be exceeded.
the media-sharing hypothesis are reinforced by the home-game-substitute hypothesis: the NFL should have the least winning-elastic revenues (equilibrium A in Figure 2), MLB should have the most winning-elastic revenues (equilibrium B in Figure 2), and the NBA equilibrium should fall somewhere between A and B. Based solely on the beta-factor elasticity propositions, one would predict that MLB would have the highest competitive imbalance and player costs, the NFL would have the lowest competitive imbalance and player costs, and the NBA would fall between the other leagues in terms of both competitive imbalance and player costs.

The sharing of winning-elastic revenue sources should have little impact on competitive balance, but it should compress player costs. The gate sharing arrangement in MLB is different in the American League, 80 percent home and 20 percent visitor, from that in the National League, approximately 90 percent home and 10 percent visitor. This compares with the 60-40 percent home-visitor formula in the NFL and the fact that gate revenues are not shared in the NBA. The elastic revenue sharing hypothesis predicts that player cost will be the highest in the NBA, next highest in MLB and lowest in the NFL. The hypothetical revenue maximization solutions of the NBA, MLB, and NFL are shown as positions A, B, and C in Figure 4.

The unique institutional arrangements of the three leagues lead to the following consensus. Due to its payroll cap and moderate revenue sharing, the NBA should have moderate player costs and the least competitive balance of the leagues. Because of the extensive revenue sharing among its teams, the NFL should have the greatest competitive balance and the lowest player costs among the leagues. As the result of the interaction of its relatively high revenue elasticity of winning with the increasing marginal cost of talent under unrestricted free agency, MLB should have moderate competitive balance with the highest player costs of the three leagues.

V. The Anomaly of the Big Show

A comparative analysis of competitive balance among the leagues employs a method used by Scully [26] for MLB and QF [24] for all three leagues. This measure of competitive balance compares the actual standard deviation of winning percentages in the league in a regular season with an idealized or fair standard deviation, which controls the number of games in a season. A competitively balanced league of teams with equal playing strengths will have a mean winning percentage of \( .500 \) and a standard deviation for the \( n \) game season of \( .5/\sqrt{n} \). The idealized standard deviation of a competitively balanced league obviously will be smaller for the leagues with the longer seasons. For example, MLB's 162 game season would imply that MLB was balanced if its standard deviation in winning percentages was .039; the 82 game NBA season would be competitively balanced at .055 and the 16 game NFL season would be competitively balanced with a standard deviation of .125 after 1977.\(^4\) The measure of competitive imbalance is the ratio of a league's actual standard deviations for each regular season to this idealized standard deviation of a "fair" league. If the ratio is one, the league is balanced competitively, and the greater the competitive balance ratio, the more imbalanced the competition within the league. The competitive balance ratios for the American (AL) and National Leagues (NL) of MLB, the NBA, and the NFL are presented in Table II.\(^5\)

Although there are several league-specific ecological factors that might affect competitive

\(^{24}\) The NFL expanded its season from fourteen to sixteen games in 1978.

\(^{25}\) MLB is split into AL and NL because there is no interleague play during the regular season.
Table II. Competitive Balance Ratios for Major Sports Leagues: 1970–92

<table>
<thead>
<tr>
<th>YEAR</th>
<th>AL</th>
<th>NL</th>
<th>NBA</th>
<th>NFL</th>
</tr>
</thead>
<tbody>
<tr>
<td>1970</td>
<td>2.410**</td>
<td>1.564**</td>
<td>2.182***</td>
<td>1.537m</td>
</tr>
<tr>
<td>1971</td>
<td>2.128</td>
<td>1.641</td>
<td>2.655°</td>
<td>1.403</td>
</tr>
<tr>
<td>1972</td>
<td>1.744x</td>
<td>2.103x</td>
<td>3.400°</td>
<td>1.701</td>
</tr>
<tr>
<td>1973</td>
<td>1.718</td>
<td>1.641</td>
<td>3.600°</td>
<td>1.724</td>
</tr>
<tr>
<td>1974</td>
<td>1.154</td>
<td>1.974</td>
<td>2.418***</td>
<td>1.478°</td>
</tr>
<tr>
<td>1975</td>
<td>1.897</td>
<td>1.897</td>
<td>2.145°</td>
<td>1.873°</td>
</tr>
<tr>
<td>1976</td>
<td>1.564x</td>
<td>2.051x</td>
<td>1.909°*</td>
<td>1.873**</td>
</tr>
<tr>
<td>1977</td>
<td>2.513**†</td>
<td>2.051†</td>
<td>1.782†m</td>
<td>1.590</td>
</tr>
<tr>
<td>1978</td>
<td>2.231</td>
<td>1.615</td>
<td>2.018</td>
<td>1.376†#</td>
</tr>
<tr>
<td>1979</td>
<td>2.333</td>
<td>1.821</td>
<td>1.873</td>
<td>1.392</td>
</tr>
<tr>
<td>1980</td>
<td>2.051</td>
<td>1.590</td>
<td>2.764*</td>
<td>1.496</td>
</tr>
<tr>
<td>1981</td>
<td>1.949xx</td>
<td>2.179xx</td>
<td>2.927</td>
<td>1.392</td>
</tr>
<tr>
<td>1982</td>
<td>1.769</td>
<td>1.590</td>
<td>2.782</td>
<td>1.704xx</td>
</tr>
<tr>
<td>1983</td>
<td>1.872</td>
<td>1.564</td>
<td>2.927**</td>
<td>1.392°</td>
</tr>
<tr>
<td>1984</td>
<td>1.487</td>
<td>1.385</td>
<td>2.091^</td>
<td>1.680°</td>
</tr>
<tr>
<td>1985</td>
<td>1.872</td>
<td>2.231</td>
<td>2.655°</td>
<td>1.568°</td>
</tr>
<tr>
<td>1986</td>
<td>1.410‡</td>
<td>1.923‡</td>
<td>2.618</td>
<td>1.672</td>
</tr>
<tr>
<td>1987</td>
<td>1.641‡</td>
<td>1.513‡</td>
<td>2.800</td>
<td>1.400x</td>
</tr>
<tr>
<td>1988</td>
<td>1.949‡</td>
<td>1.949‡</td>
<td>2.873**</td>
<td>1.344</td>
</tr>
<tr>
<td>1989</td>
<td>1.667</td>
<td>1.513</td>
<td>2.945**</td>
<td>1.440‡</td>
</tr>
<tr>
<td>1990</td>
<td>1.462x</td>
<td>1.462x</td>
<td>3.164</td>
<td>1.624</td>
</tr>
<tr>
<td>1991</td>
<td>1.564</td>
<td>1.564</td>
<td>2.873</td>
<td>1.744</td>
</tr>
<tr>
<td>1992</td>
<td>1.615</td>
<td>1.692</td>
<td>2.891</td>
<td>1.656</td>
</tr>
</tbody>
</table>

* Number of expansion teams (2 NL and 2 AL expansion teams in 1969 are shown as 1970).
† Implementation of free agency agreements.
° Competition from other professional leagues.
‡ Collusion by MLB owners against free agent signings.
m Merger with competing leagues.
^ Implementation of salary cap.
x Strike or lock out seasons; xx over 50 days.
# NFL season was lengthened from 14 to 16 games and “balanced scheduling” began in 1978.
Sources: [27; 28; 29]

Balance, some relevant intraleague and interleague observations can be made. First, in the era of free agency, there does not seem to be a worsening of competitive balance in MLB; on the contrary, both leagues are becoming increasingly competitive. This is consistent with the findings of Scully [26] and QF [24]. Second, the NBA is clearly the least balanced of all three leagues and there is no tendency for this balance to be increased by the imposition of the salary cap in 1984. This is consistent with the findings of Noll [19] and QF [24]. Third, the NFL is clearly the most balanced of the three leagues and there does not seem to be any change in that balance over the period, including the rival interleague competition in 1983–85. This is consistent with the ecological differences inherent in the games could lead to variation in competitive balance that could make interleague comparisons problematic. For example, there are twelve players on the active NBA team roster, twenty-five for MLB and forty-five for NFL teams. Competition variances may reflect the greater importance of a few players for team production on smaller rosters. This hypothesis would also be consistent with the relative balance in the NFL and the imbalance in the NBA.

It is acknowledged that this was a period of continuous expansion for the NBA, and that the influx of newer noncompetitive teams and the resulting talent dilution may be a source of imbalance.

There does seem to be an increase in balance with the implementation of balanced scheduling in the NFL after 1977. Competition from the USFL did lead to higher salaries for NFL players.
Table III. Parameter Estimates for the General Model of Professional Sports Leagues: 1990–92

<table>
<thead>
<tr>
<th></th>
<th>WIN3</th>
<th>POP</th>
<th>YEAR</th>
<th>NEW</th>
<th>TVDEAL</th>
<th>XPN</th>
<th>CON</th>
<th>$R^2$</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>REVENUE</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MLB</td>
<td>0.599</td>
<td>0.276</td>
<td>0.074</td>
<td>0.392</td>
<td>3.97</td>
<td>0.729</td>
<td>78</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(3.53)</td>
<td>(6.00)</td>
<td>(6.56)</td>
<td>(7.72)</td>
<td>(31.37)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NBA</td>
<td>0.488</td>
<td>0.177</td>
<td>0.162</td>
<td>0.182</td>
<td>0.323</td>
<td>3.19</td>
<td>0.901</td>
<td>81</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(7.14)</td>
<td>(6.81)</td>
<td>(7.86)</td>
<td>(2.21)</td>
<td>(5.94)</td>
<td>(3.69)</td>
<td>(39.28)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NFL</td>
<td>0.120</td>
<td>0.032</td>
<td>0.075</td>
<td>0.034</td>
<td>3.88</td>
<td>0.904</td>
<td>84</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(9.98)</td>
<td>(4.04)</td>
<td>(17.09)</td>
<td>(3.22)</td>
<td>(86.48)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EXPENSES</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MLB</td>
<td>1.304**</td>
<td>0.223</td>
<td>0.140</td>
<td>4.39</td>
<td>0.852</td>
<td>78</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(10.78)</td>
<td>(6.18)</td>
<td>(10.14)</td>
<td>(42.14)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NBA</td>
<td>0.261</td>
<td>0.049</td>
<td>0.266</td>
<td>0.027*</td>
<td>3.05</td>
<td>0.883</td>
<td>81</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(5.82)</td>
<td>(2.66)</td>
<td>(21.71)</td>
<td>(5.8)</td>
<td>(60.69)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NFL</td>
<td>0.089</td>
<td>0.011*</td>
<td>0.135</td>
<td>3.69</td>
<td>0.828</td>
<td>84</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(3.48)</td>
<td>(1.05)</td>
<td>(19.50)</td>
<td>(159.0)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Numbers in parentheses are "t" ratios.
*Not significantly different than zero @ .01.
**Significantly greater than unity @ .01 for log coefficients.

findings of QF [24]. Clearly the NFL is the most competitively balanced, MLB is slightly less competitively balanced, and the NBA is the least competitively balanced of the three leagues. These estimates of competitive balance, which are corroborated by every recent study, are consistent with the hypotheses of the general theory of sports leagues developed above.

Because of severe data limitations and the possible impact of collusion on the competitive balance of MLB in the period immediately preceding the collusion settlement in 1990, empirical investigation of the revenue parameters in the model used above is confined to the post-collusion period 1990–92.29 Consistent and reliable data for all three major sports leagues are available for the estimates of both revenue and cost equations from a recent annual series of three reports by Financial World [1; 20; 21].30 The parameters of the revenue equation of the general model can be estimated using the double-log specification:

$$\ln R_{it} = \ln R_0 + \alpha \ln p_{it} + \beta \ln w_{it} + \rho t_{it} + e_{it}$$

where $\rho$ is the exogenous revenue growth for team $i$ over the $t$ years of the survey, $p_{it}$ is the population in 1990 of the home market SMSA (POP) and $w_{it}$ is the average winning percentage for the current and two previous seasons (WIN3) for each team.

The results are presented in Table III for each of the three leagues.31 The estimates of the

29. The collusion of the MLB owners was ruled in arbitration in 1987, 1988 and 1989. "Collusion I, II and III" was resolved with a $280 million settlement for players in 1990.
30. Although most financial data of the clubs are confidential, these estimates are consistent with what limited information is available. Noll has consistently testified to the unreliability of financial data released by team owners. The 1990 FW data are used by QF [24] and Zimbalist [30].
revenue elasticity of winning are consistent with the hypotheses of the general theory of sports leagues. MLB has the highest winning elasticity (β = .6), the NBA has less (β = .5), while the NFL has the lowest revenue elasticity of winning (β = .12). These results are precisely what should be expected under the current institutional configuration of the three leagues.\footnote{32} The large market revenue advantage is the greatest in MLB, less for the NBA, and the least for the NFL. Finally, annual total revenue growth has been the highest in the NBA at 16 percent over the three-year period, compared to growth of 7.5 percent in both MLB and the NFL.

The parameters of the cost equations can be estimated from the double-log specification:

$$\ln C_{it} = \ln C_0 + \gamma \ln p_{it} + \delta \ln w_{it} + \mu t_{it} + \epsilon_{it}$$

where $\mu$ is the rate of increase in expenses for team $i$ over the three years of the survey. The estimates of these parameters are also presented in Table III. As expected, the cost elasticity of winning (\(\delta\)) is significantly greater than unity for only MLB. This implies that the nondecreasing cost conditions required for profit maximization exist only in MLB. Furthermore, the estimates of $\gamma$ indicate that the only appreciable diseconomies of market size exist in MLB. This supports the contention that there are diminishing marginal returns to talent in MLB and that the large market teams are paying significantly more than the small market teams for the same level of talent in the post-collusion period.

The competitive balance of a simulated league consisting of an average large market team, and an average small market team can be determined from the parameter estimates in Table III and the average populations for large and small markets in each league. The average population for the larger half of MLB home markets is 5.61 million compared to 2.31 million for the smaller half; in the NBA the larger half has an average SMSA population of 5.58 million and 1.67 million

\footnote{32 Based on attendance figures, ticket prices and various characteristics about the teams and their respective home markets for the 1970 and 1971 seasons, Noll [18] concluded that market size (\(\alpha\)) was most important in MLB, less important in the NFL and least important in the NBA. In contrast, he observed that winning, as a determinant of home attendance, was “extremely important” in the NBA (\(\beta = .8\)), moderately important in the NFL (\(\beta = .6\)), and least important in MLB (\(\beta = .4\)).}
for the smaller half; and in the NFL the larger half of the home markets is 5.81 million and the smaller markets averaged 1990 populations of 1.95 million. Substitution of the estimates of $\alpha$ and $\beta$, and the average populations for small and large markets into the revenue maximization solution (2) yields a simulated large/small competitive balance of .603/.397 in the NBA, .510/.490 in the NFL and .648/.352 in MLB for the revenue model. These solutions are shown graphically in Figures 5, 6 and 7. Because of the institutional constraints of the salary cap in the NBA and the illusion of conditional free agency in the NFL, the cost parameters have been abrogated and the revenue model should yield accurate predictions for these two leagues. The estimated impact of free agency in MLB, however, is biased. The predicted competitive imbalance for MLB is exaggerated, because the naive revenue maximization model (2) omits the cost parameters of the marginal cost of talent ($\delta$) and (dis)economies of market size ($\gamma$).
Figure 8. Simulation of Major League Baseball under Free Agency

An unbiased competitive balance solution of the general model (1) for MLB yields the corrected large/small competitive balance of .517/.483 for MLB as shown in Figure 8. In the NFL and the NBA profit maximization for the individual teams is equivalent to the collusive revenue maximization of the league. In contrast, the existence of free agency in MLB generates a profit maximization for the individual teams that yields a competitive balance greater than that required for revenue maximization. In this regard, the general theory, in which the NBA and the NFL are reduced to special revenue maximization cases, yields competitive balance predictions that are consistent with the actual competitive balance of all three major sports leagues.

VI. Conclusion

The general theory developed in this paper leads to three major conclusions concerning free agency, revenue sharing and the payroll cap. First, unrestricted free agency during the post-collusion era in MLB has allowed the individual team to retain its identity as the firm, and MLB competition has remained relatively balanced in spite of a clear revenue advantage for the large market teams. Specifically, the increasing marginal cost of talent serves to mitigate on-field reflections of the revenue disparity among teams. The general theory predicts that MLB would have moderate competitive balance with the lowest player exploitation of the three leagues. Second, the sharing of winning-elastic revenues among teams paradoxically does not affect competitive balance among teams, but because it adversely affects the incentives of teams to win, it leads to the exploitation of players. If revenue sharing is to increase competitive balance, it is necessary that the revenue be winning inelastic. Consequently, the extensive sharing of winning-elastic revenue in the NFL increases competitive balance to virtual parity, whereas the extensive sharing of winning-elastic revenue in the NFL leads to the depression of player salaries. According to

33. The actual large/small mean winning percentages for the three seasons of the study were .492/.508 in MLB (.504/.496 in AL and .478/.522 in NL); .515/.485 in the NBA; and .537/.463 in the NFL. The simultaneous profit maximization simulation for MLB requires that $C_0$ be treated as a variable portion of revenue.
the general theory, the NFL is predictably the most competitively balanced, but also the most exploitive of the leagues. A third conclusion of this paper is that the payroll cap is a unique form of cost-sharing collusion, and that, because of its implementation, the NBA is a cartel of teams acting as a single firm. If NBA teams collusively behave as the firm, then profit maximization is reduced to revenue maximization for the league. The salary cap and the cost-sharing collusion of the NBA predictably lead to the least competitive balance of the three leagues. The policy implications of the general theory are immediate. The imposition of a payroll cap allows a cartel of teams to collusively behave as the firm, and the capping of team payrolls leads to the increased exploitation of players and decreased competitive balance within the league.

This reconsideration of the economics of professional sports ultimately leads to the questioning of “competitive balance” as a meaningful objective of professional sports leagues. The inelastic-revenue sharing of the NFL creates a balanced economic field on which to play, and “on any given Sunday any NFL team can defeat any other.” But is the NFL an example of successful parity-engineering or has it just become a victim of self-imposed mediocrity? Critics of free agency argue that competitive balance may be more the result of dragging a good team down than bringing a bad team into contention. This implies that free agency has a negative sum effect on the quality of talent within the league, and that “competitive balance” is the result of two equally bad teams beating each other. The importance of team specificity of talent in professional sports and the existence of information asymmetries inherent in the acquisition of players in the external market can create a world of positive transactions costs, in which this criticism would be valid.

References