Scottish Journal of Political Economy, Vol. 47, No. 4, September 2000

THE ECONOMICS OF AMERICAN SPORTS LEAGUES

John Vrooman[*]

Abstract

Player mobility in North American sports leagues is limited by labour market constraints designed ostensibly to allow teams to recoup player development costs and to maintain competitive balance within leagues. Theory holds that the distribution of talent will be invariant under any institutional configuration, and that while rules to constrain movement will serve to enhance monopsonistic exploitation of talent, they will have no effect on competitive balance. This paper reprises a general theory through which the effects of the labour market constraints can be comparatively analysed for American sports leagues. The economics of sports has been relegated to the realm of labour theory by the unnecessarily limiting assumption that owners of sports clubs are single-minded profit maximisers. This paper also presents a theory that seeks to unify capital market decisions of financial leverage and ownership syndication with operational labour market decisions for athletic talent and an owner's desire to win.

> The financial results of the last season prove that salaries must come down. We believe that players insisting on exorbitant prices are injuring their own interest by forcing out of existence clubs which cannot be run and pay large salaries except at a large personal loss ... In view of these facts, measures have been taken by this league to remedy the evil.
>
> —The National League of Professional Baseball Clubs, following its adoption of the reserve rule in 1879

I Introduction

The lone professional sports league in America at the dawn of the twentieth century was a pale precursor of what was to come by the dawn of the twenty-first. A beleaguered National League of Professional Baseball Clubs (NL) monopoly, which had survived challenges from rival leagues over two decades, was again being faced with the formation of yet another rival in the

[*] Vanderbilt University

American League (AL) of 1900. This rival AL would combine successful minor-league cities with abandoned NL markets, and raid NL player rosters by offering salaries higher than the $2500 NL salary cap. After three successful AL seasons, an embattled NL sued its junior rival for peace and negotiated the National Agreement of 1903. Under this agreement, the NL and AL would operate as separate but equal leagues, unified by common playing rules and mutually recognized player contracts and territorial monopolies. The resulting dual-league monopoly enjoyed a remarkable franchise stability over the next half century. Baseball players did not fully share in the prosperity, however, because it was largely derived from the leagues' monopsony power over players exercised through the century-old *measure* of the reserve clause, which bound the rights of reserved players to one team for life.

The economics of American sports leagues finds its origin in a seminal paper by Rottenberg (1956) in which the economic justification for the reserve clause is blown away. In this prescient piece Rottenberg anticipates almost every subsequent labour market restriction in addition to the reserve clause, and presages the important parameters of free-agency debate. Subsequent analysis over the last three decades has followed three academic avenues, each headed by economists who have significantly extended economics into professional sports in their own way. The mathematical foundations for the economics of sports are the results of pioneering work by Quirk and coauthors (El Hodiri and Quirk, 1971; Fort and Quirk, 1995; Quirk and El Hodiri, 1974; Quirk and Fort, 1992); empirical investigations of the subtleties of the players' labour markets follow the lead of Scully (1974, 1989, 1995); and policy avenues are explored by Noll (1974, 1991, 1995) and Noll and Zimbalist (1997). This paper is a reprise of a recent trilogy of articles (Vrooman, 1995, 1996, 1997a) originally written as critical syntheses of these earlier works.

II The Players' Labour Market

The invariance proposition

Unfettered free agency for the players in North American professional sports leagues is an illusion. While the evolution of the legal foundations for free agency spans two decades from the 1972 *Flood v. Kuhn* supreme court decision[1]

[1] *Flood v. Kuhn*, 107, U.S. 258 (1972) 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 through the reserve clause. It concerned the continuation of an antitrust exemption enjoyed by MLB from 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). Chief Justice Burger admitted that the Court was originally in error, but that Congress should correct the mistake. The free agency issue was resolved in *The Collective Bargaining Agreement of 1976*, in which the limits of the reserve clause were reduced to six years of MLB service.

in Major League Baseball (MLB) to the 1992 *McNeil, et al. v. NFL* decision[2] in the National Football League (NFL), the coevolution of ancillary labour market constraints has continued to severely limit player mobility in all leagues. 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% of NBA designated gross revenues (DGR).[3] The ostensible purpose of the NBA payroll or 'salary' cap was to 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.[4] Although free agency eligibility requirements have been lenient in the NFL, strict compensation rules for teams losing free agents had thwarted player movement until *McNeil*.[5] In spite of free agency concessions made by NFL owners in *The Collective Bargaining Agreement of 1993*, movement of players among teams continues to be limited after *McNeil*, because of the joint imposition of a payroll cap at 64% of league-wide DGR. While *nominal* free agency in MLB finds its genesis in the *Collective Bargaining Agreement of 1976*, the series of subsequent *Bargaining Agreements* that winds through the free agency era has continued to subject players with less than six years MLB service

[2] 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 with 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).

[3] At the same time that Flood initiated action against MLB, Oscar Robertson, a player for the Milwaukee Bucks, filed a class action suit in 1970 against the reserve clause of the NBA: *Robertson v. National Basketball Association*, 369 F. Supp. 867 (1978). The free agency issues of the Robertson case were resolved in an out-of-court settlement, culminating in the *Collective Bargaining Agreement of 1976*. The eligibility requirement for the NBA has 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.

[4] NBA clubs may exceed the cap for a variety of grandfather or 'Larry Bird' exemptions. Teams may exceed the 'soft' cap to resign their own free agents or to extend contracts of their veterans. Teams that are 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% of the departed player's salary. The NFL cap is a hard cap in that in cannot be exceeded.

[5] From 1977 to 1993, the NFL operated under the 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'. Free agency for NFL players during this period was nominal because of the extremely severe conditions of compensation to the original team. 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.

to the binding constraints of the reserve clause.[6] Only those players with six or more years MLB service (approximately one-third of MLB players) are eligible for free agency.

The arguments for payroll caps and other free agency restrictions under modified reserve are the same as those traditionally made for the strict player reserve system. It is usually held that player salaries should be controlled initially so that teams could recoup their investment in player development. Also, it is maintained that talent will gravitate toward large markets under unbridled free agency, and that constraints are necessary to defy gravity and maintain competitive balance within a league. Conventional economic theory lends qualified support to the first of these justifications, but stands in diametric opposition to the second (Vrooman, 1996). Theory has steadfastly 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 changes. 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 (Rottenberg, 1956, p. 255),

and used by Demsetz as an example of the 'Coase theorem' (1960) 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 (Demsetz, 1972, p. 17).

If the distribution of talent is invariant, then institutional restrictions serve only to redistribute income from players to team owners. The purpose of this section is to review the implications of economic theory within the institutional configurations of professional sports leagues as they have evolved.

General theory

The economics of professional sports leagues 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

[6] Before the *Bargaining Agreement of 1976*, the reserve clause of the standard players contract had bound players to their original team for perpetuity. Since 1976, eligibility requirements for free agency have remained the same. Any player with 6 or more years in Major League Service who has not executed a contract for the next succeeding season shall be eligible to become a free agent. Major changes since then have involved the type of compensation (draft picks or players) due to the team that loses the free agent.

institutional constraints.[7] Rottenberg's *invariance proposition* was formalized by Quirk and El-Hodiri (QE) (1971, 1974) where:

> 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 (1974, p. 58)

or more simply:

> big cities have winning teams and small cities have losing teams (1974, p. 45).

The competitive equilibrium solution for a simplified two-team league is shown in Figure 1 for a large market team 1 (left to right) facing a small market team 2 (right to left). The marginal revenue functions show the superior revenue potential of the large market club such that $MR_1 > MR_2$ at ·500. If the talent of players is not team specific, and if the cost per unit of talent is the same between clubs ($MC_1 = MC_2$ at ·500),[8] then decreasing marginal productivity of talent implies that marginal cost functions are increasing and symmetrical around ·500, similar to those in Figure 1. Profit maximization yields A_1 and A_2 for the respective clubs, and the revenue advantage for by team 1 results in its $w_1/(1-w_1)$ dominance of team 2.

To see why this is a stable equilibrium, consider competitive outcomes other than w_1. For any outcome less than w_1 for team 1, it is profitable for team 1 to acquire talent from its opposition. For any outcome greater than w_1 for team 1, it is profitable for team 2 to acquire talent from team 1. If the league finds itself at *B*, for example, then team 1 will seek to win more games (team 2, less), and if the league finds itself at *C*, team 1 will seek to win less games (team 2, more). Profits for both clubs in the league are simultaneously maximized at $A_1 - A_2$. The uniqueness of this equilibrium depends on the zero-sum restrictions of the league (if a team wins, its opponent must lose), and the equilibrating mechanism is the cost per unit of talent. Zero-sum league equilibrium requires $\Sigma w_i = n/2$ for the *i*th team in a *n*-team league. If a league is initially supercompetitive such that

[7] In Rottenberg's analysis, the most unambiguous way of enhancing competitive balance is to 'break up the Yankees' by dividing the large monopoly markets until markets are about the same size. 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' (1992, p. 14). This is probably not true. Most economists see the invariance proposition, like the Coase theorem, as a limiting case.

[8] The equal cost per unit of talent assumption is made by several authors (Demmert, 1973; El Hodiri and Quirk, 1971; Fort and Quirk, 1995; Quirk and El Hodiri, 1974; Quirk and Fort, 1992; Vrooman, 1996). Oi suggests 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' (Quirk and El Hodiri, 1974, p. 66n). Demmert notes that 'the empirical question of what particular market variations affect a club's costs of acquiring athletic talent is of some interest' (1973, p. 51). Commercial opportunities in larger markets supplement players' incomes and reduce both the wage rate and the marginal costs for teams in large markets. Although endorsement opportunities create positive externalities of market size, recent difficulties in the free agent signings of large market franchises may suggest *dis*economies of market size.

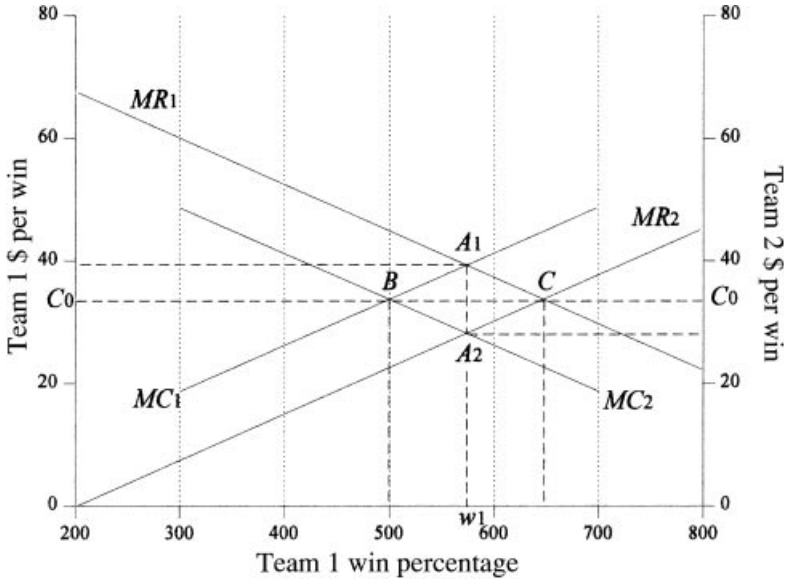

Figure 1. League equilibrium.

$\Sigma w_i > n/2$, then an excess demand for talent will cause the cost per unit of talent to rise and both *MC* curves to be vertically displaced upward. Talent acquisitions and wins for both teams will decline until the zero-sum restrictions hold. If a league is initially subcompetitive such that $\Sigma w_i < n/2$, then the cost per unit of talent will fall, both *MC* curves shift downward. Talent acquisitions and wins will increase for both clubs until $\Sigma w_i = n/2$.

Quirk and Fort (QF) base their rendition of Quirk's earlier work in 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) (Quirk and Fort, 1992, p. 273).

This assumption implies that marginal costs are the same for all teams regardless of its winning percentage.[9] QF's model amounts to a special revenue

[9] Hunt and Lewis (HL) assume that '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' (1976, p. 940). The marginal cost curve for team 1 is 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 at *C* in Figure 1. Although this is a slightly different argument than that of QF, it yields the same competitive equilibrium. Like the argument of QF, HL's position is a special case, and unnecessarily limits the inquiry into the effects of free agency on competitive balance. 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' (1974, p. 38).

maximization case of the general model presented here. As a limiting case, the marginal cost assumption of QF yields a revenue maximization solution at C in Figure 1. Clearly, the omission of the increasing marginal cost of winning (decreasing marginal productivity of talent) overstates the competitive imbalance implied by revenue disparity.

The Yankee paradox

The customary argument for balanced competition is couched in terms of league-welfare optimality. It is argued that the product of professional sports teams is the game that is jointly produced between teams, and that the quality of the games is determined by the uncertainty of the outcomes of games between members of the league. The objectives of the teams in the league are necessarily interdependent, because each game requires the existence and zero-sum performance of two teams. In this peculiar case, the accumulation of talent in the singular pursuit of maximum profit by individual clubs may lead to significant negative externalities and a self-defeating dominance of the league by large-market clubs (the *Yankee paradox*). Quirk and Fort concur that:

> one of the key ingredients of the demand by fans for team sports is the excitement generated because of the uncertainty of outcome of league games (1992, p. 243).

The argument follows a second proposition by Rottenberg that the interdependence of profits among teams will be internalized by the large market owners.

> The wealthy teams will usually prefer wining to losing. If they do they will prefer winning by close margins to winning by wide ones. If their market behavior is consistent with this objective ... playing talent will be more or less equally distributed among teams (Rottenberg, 1956, p. 255).

Competitive balance is thereby seen as being superior *a priori* to large market dominance in terms of league welfare. The small market clubs will be more profitable if they have a fair chance to win, and the large market clubs will be more profitable because they provide a game with an uncertain outcome.[10] While first part of this argument is true *a priori*, the second part is not so easily demonstrated. Whether large market profits are further enhanced through league competitive balance or the closeness of the competition remains an empirical question.

[10] Neale labelled the negative interdependence the 'Louis-Schmelling Paradox' but refers to the pre-1964 New York Yankees (as does Rottenberg) in making the observation: 'the greater the economic collusion and the more the sporting competition the greater the profits' (1964, p. 2).

Empirical evidence for the *Yankee paradox* is mixed.[11] Demmert tests the proposition and concludes:

> Rottenberg bases much of his argument on the maintained hypothesis that the returns to narrow victory are larger. Such a hypothesis does not appear to merit support in light of our estimates ... the incentive of the individual club is to win, and not necessarily by a close margin (Demmert, 1973, p. 67).

In an inquiry into a league-optimal dominance, Hunt and Lewis conclude that:

> revenue is maximized for the division when the large SMSA team wins the division 43% of the time (1976, p. 940).[12]

Canes argues that:

> a policy aimed at equalizing team strengths would transfer quality from areas where winning is more valued to areas where it is less valued ... Since those who gain place lower value on increased wins than those who lose, *net* social benefits would be reduced. This raises doubts whether equality of team strengths is a desirable goal (1974, p. 82).

These results suggest that competitive balance may not be consistent with the joint profit maximization of the teams in the league, and that it may not be in the best interest of large market clubs to fully internalize the Yankee paradox.

Quality of the game

A more convincing argument for competitive balance concerns the *a priori* proposition that evenly matched playing strengths among clubs are necessary to create the combination of league talent that yields the maximum league product. If this is true, then the talent allocation consistent with large market dominance produces contests that are inferior with respect the league's total product. It is profitable for large revenue clubs to acquire players who may have a higher marginal *revenue* product, but who still may have a lower marginal *physical* product than they would with small revenue teams. As a result, the league-wide talent combination under large market dominance is necessarily submaximal in terms of total quality. This leads to the conclusion that the distribution of talent that is consistent with the maximum league product occurs when the marginal physical product of talent is equivalent among clubs in the league. The accumulation of physically redundant talent by large market clubs, therefore, leads to an inferior, submaximal league product.

[11] In Noll's study of the 1970–71 seasons a close pennant race and games behind variables are not significant factors in explaining MLB attendance, but 'recent pennant win' is. 'This suggests that aggregate league attendance will be substantially higher if several teams alternate in winning pennants than if one team tends to dominate' (1974, p. 123).

[12] Hunt and Lewis (1976) base their estimation of a league-optimal level of competitive dominance of the large market team on a comparison of revenue maximization for the individual team (winning the division 80% of the time) compared to the revenue maximization for the league (winning the division 43% of the time).

The proposition is also demonstrated in Figure 1. Any team playing above ·500 encounters a diminishing marginal productivity of talent and a higher marginal cost of talent than the team it dominates. Under these conditions, the maximum league product is consistent with a competitively balanced league at B. A comparison of the large market dominance equilibrium $A_1 - A_2$ with the league maximum talent production at B shows the reduction in the league quality. A redistribution of talent from team 1 to team 2 in the pursuit of competitive balance (w_1 to B) would improve the quality of team 2 more than it would harm team 1. League costs would be minimized and net gains to the total league product would be exhausted at the optimal talent distribution B, but the profits of both teams would be submaximal. Conversely, a profit maximization solution of large market dominance at $A_1 - A_2$ necessarily implies a suboptimal combination of league talent.

Revenue sharing paradox

There is major confusion in the literature about the impact of revenue sharing on competitive balance.[13] On one side, QE and QF propose that revenue sharing will have little impact on competitive balance, because of the disincentive effects caused by the unique interdependence among the wins by teams in a league. On the other, Scully suggests 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 (1989, p. 80).

Although the second proposition is intuitively appealing, it is not necessarily true. Scully's model allows for the increasing marginal cost of talent, but it overlooks the zero-sum restrictions of the league $\Sigma w_i = n/2$. As a result, the model does not have a unique competitive solution, and while it does yield accurate descriptions of large market dominance, the omission leads to an oversight of the disincentive effects of the negative interdependence on competitive balance.

The zero-sum restrictions of the league $\Sigma w_i = n/2$ imply negative interdependence between any team's fortunes and those of the rest of the league: $\partial w_2/\partial \mathrm{w}_1 = -1$ for a simple two-team league. Under a revenue sharing formula where α is the home team's revenue share and $(1-\alpha)$ is the visitors share, the internalization of the interdependence alters the marginal revenue function of team 1 *with respect to* w_1 such that:

$$\begin{aligned} MR_1^* &= \alpha \partial R_1/\partial w_1 + (1-\alpha)\partial R_2/\partial w_2 \; \partial w_2/\partial w_1 \\ &= \alpha \partial R_1/\partial w_1 - (1-\alpha)\partial R_2/\partial w_2. \end{aligned}$$

[13] The three major sports leagues provide an excellent controlled environment for the effects of revenue sharing. The NFL shares gate revenue 60/40 (home/visitor) and divides almost all other revenue equally. MLB shares the gate 80/20 and in the NBA (also in the National Hockey League) home teams do not share gate revenue with the visitors.

If marginal revenue functions are equally adjusted for other teams in the league (team 2), then simultaneous profit maximization across the league *with* revenue sharing $MR_1^* - MC_1 = MR_2^* - MC_2 = 0$ yields the same competitive balance as profit maximization *without* revenue sharing $MR_1 - MC_1 = MR_2 - MC_2$. 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 directly with its own ability to lose as a visitor. With revenue sharing, it pays a team to win at home, and lose (so that its opponent can win) on the road. Thus, while gate sharing does not necessarily make the marginal revenue functions of the teams identical, as intuition would suggest, it does adversely affect the winning incentives of the teams in proportionately the same way.[14]

Revenue sharing displaces the revenue functions of both clubs proportionately downward in Figure 1.[15] These equal revenue disincentives create a subcompetitive league $\Sigma w_i < n/2$ and downward pressure on the cost per unit of talent for both clubs. Accordingly, both *MC* curves shift downward until equilibrium is restored at w_1 for a lower value of C_0 at $(2\alpha - 1)C_0$.[16] The important difference between the disincentive effects of profit maximization with or without revenue sharing lies, therefore, in the reduction of player salaries, not in the competitive balance solution. Revenue sharing results in proportionately lower marginal revenues from winning for all teams and lower player costs at equilibrium, and as the home teams α share becomes smaller, the exploitation of players increases. This leads to the simple, but important conclusion that revenue sharing, in equilibrium, will effectively increase the exploitation of players, *but it will not affect competitive balance*.

Myth of the salary cap

During the second decade of the free agency era, the payroll cap has co-evolved with free agency as its countervailing compromise. The ostensible purpose of the

[14] Noll concludes that the 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' (1991, pp. 33–4). Zimbalist (1992) makes a similar argument for MLB. It does not matter. Even under one-way revenue sharing, where the large market team retains an α share gives the rest to the small market club, the marginal revenue functions for the teams become: $MR_1^* = \alpha MR_1$ and $MR_2^* = MR_2 - (1-\alpha)MR_1$, where $MR_1^* - MC_1 = MR_2^* - MC_2$ still implies the same competitive balance as without revenue sharing: $MR_1 - MC_1 = MR_2 - MC_2$.

[15] See either QF (1995, p. 30) or Vrooman (1995) for the effects of revenue sharing on marginal revenue equilibria.

[16] If $MR_1^* - MC_1 = MR_2^* - MC_2$ implies $MR_1 - MC_1 = MR_2 - MC_2$, then revenue sharing in the above formula suggests that player costs approach $(2\alpha - 1)C_0$ in revenue-sharing equilibrium. In general equilibrium for the model, $C_0 = (MR_1 + MR_2)/2$ at ·500.

payroll cap is to foster competitive balance at B in Figure 1. QF concur 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 strength (1992, p. 287).[17]

Unfortunately, this widely held proposition is not true either.[18] Upon closer analysis, the payroll cap may not serve to improve competition, but rather to worsen it. Consider the possibility that the cap is a transparent method of controlling total player costs, and it indirectly allows for the maximization of profits *for the league as a whole*. Under a payroll cap, all teams share equally in player costs that are ultimately a proportion of combined league revenues. In this case, profits for individual teams will be maximized at the league revenue maximum shown at C in Figure 1. To see why this is true, recall the counterintuitive disincentives caused by a zero-sum game. The marginal profit function for team 1 under a salary cap arbitrarily set at a μ share of league DGR becomes:

$$\begin{aligned}\partial\pi_1/\partial w_1 &= \partial R_1/\partial w_1 - \mu(\partial R_1/\partial w_1 + \partial R_2/\partial w_2\ \partial w_2/\partial w_1)\\ &= (1-\mu)\partial R_1/\partial w_1 + \mu\partial R_2/\partial w_2\end{aligned}$$

where marginal costs become a negative μ share of marginal league revenue. Simultaneous profit maximization across teams (with team 2) yields the conditions for league revenue maximization, $\partial R_1/\partial w_1 = \partial R_2/\partial w_2$. If player costs under the cap are proportional to revenue, then each team has an identically dampened incentive to win at home and enhanced incentive to win on the road (so that its opponent will lose and reduce player costs). The imposition of a payroll cap effectively eliminates increasing marginal cost of talent from the profit maximization calculus, and thereby allows the teams of the league to collusively behave as the revenue maximizing firm. As a result, the capping of team payrolls ironically promotes competitive *imbalance* within the league. Compare the collusive payroll cap solution C with the competitive free agency solutions A_1 and A_2 in Figure 1.

The irreversibility proposition

If individual talent is team-specific to any degree (players are not interchangeable parts among teams), then the possibility arises that a blind pursuit of competitive balance *per se* may result in an inferior league product of equally

[17] QF's salary cap solution is the point where a maximum cap set below C_0 intersects MR_2. In this case the 'ideal' cap for a balanced league would be set at the value of MR_2 at ·500.

[18] It does not necessarily follow that equal spending among teams implies equal playing strengths. 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, but it is still a function of the (dis)economies of the home markets in which the teams play. Given constant player costs, talent will still migrate within the league to markets that promise the highest non-playing income possibilities or the highest non-income benefits such as a chance to play on a championship team.

bad teams beating one another. The existence of each professional sports club necessarily implies that it is a coalition of individual players for which the collective results are greater than the sum of the individual results. Professional sports teams would simply disband in the absence of residual team talent or net synergy among players. If the value of a player is partially attributable to his team, then that player's talent is team-specific. Not only is the talent of the team-specific player asymmetrical or irreversible among the teams of the league, but so is the knowledge of team-specificity. The most complete information about the team-specificity of talent resides with the owner of the team for which the player currently plays. Incomplete information is, therefore, inherent in any transaction that involves player movement from team to team. In either the acquisition of players for salaries under free agency,[19] or club-to-club compensation under the reserve system, information about the team specificity of a player's talent is *asymmetrically* incomplete.[20] The important question is whether the asymmetry of information is any different under free agency than it was under the reserve system.

Market for lemons. Lehn (1990a, 1990b) argues that, *because of free agency*, the market for MLB talent is subject to self-selection and that the 'lemons model' (Akerlof, 1970) may apply to the free agent market.

> Presumably, the club for which a player has performed has more information with which to estimate the player's future performance than do the clubs that have not employed the player ... If such asymmetry exists, then, the expected performance of free agents will be lower than the expected performance of eligible players who do not become free agents (Lehn, 1990, p. 254).

This proposition is not necessarily true, because it is the information asymmetry that creates the adverse selection of free agents on the market, *not free agency itself*. In order for free agency to affect the information asymmetry, it must be shown that any additional uncertainty about the team-specificity of talent under free agency affects the imbalance of certainty that exists between the former owner, on the one hand and the group of potential owners, on the other. If the acquisition of talent is subject to the same information asymmetries between former and prospective owners under free agency or the reserve system, then the labour market for playing talent is subject to the same degree of adverse selection under either regime.

[19] In player for player trades information is incomplete about the team specificity of talent, but the uncertainty is symmetrical between trading teams. In trades involving many for few players there is an asymmetry of risk. Large market teams often acquire a single talented player for several less talented players (the sum of whose talent may be less than the single player) in a trade with a small market team. Although the expected value of the single player is higher, the variance of the performance of the several players is lower unless their performance is perfectly correlated.

[20] In a Coasian sense, talent and information asymmetries create marginal product and marginal cost asymmetries in Figure 1 between those clubs who lose players and those who acquire them.

Winner's curse. In their analysis of the inefficiencies of the auction mechanism in the free agent market Cassing and Douglas (1980) conclude that free agents 'really are paid too much'. 'They argue that the acquisition of free agent talent is subject to a winner's curse:

> If a team bids on, say, ten free agents, then there is a tendency to win a 'biased' set of players—those for which the bidder has overestimated value. Increasing the number of bidders only increases the bias effect (Cassing and Douglas, 1980, p. 112).

The bids by potential new owners may on average estimate the 'true' worth of the player, because of information asymmetries and uncertainty, but again, this is not necessarily attributable to free agency. The asymmetry of information leads to an overpayment of free agent players, but only if there are relatively few eligible players and several prospective bidders. In the case of the winner's curse, the information asymmetry that exists under either free agency or the reserve system is a necessary, but not sufficient condition for the overbidding. Information asymmetries create the bias only because the auction process exists in a monopoly environment. Therefore, it is the arbitrary limitation of the number of free agents, not necessarily free agency, that augments the bias of the winner's curse.

Shirking. Although the information asymmetries are similar under the reserve system and free agency, the free agency transaction may still be unique in one important respect. A free agency contract usually involves an owner and a player in a one-time deal, whereas the contracting parties in reserve transactions are members of an ongoing league. The free agent transaction, therefore, may be uniquely inefficient in that it is not self-enforcing and is, therefore, subject to post-contractual opportunism. Lehn (1990a, 1990b) finds that players who were eligible for free agency and signed with new teams were more likely to experience post-contractual injury and lost playing time than those players who were resigned by their own team.[21] Lehn suggests that the length of the free agent contracts would become shorter as owners try to minimize risk of losses due to opportunistic behaviour.

Empirical evidence

On the surface, evidence on the effects of free agency on competitive balance seems contradictory. Early studies (Cymrot, 1983; Daly and Moore, 1981) have found that free agents migrate from winning teams in small markets to large markets, while later studies (Balfour and Porter, 1991; Quirk and Fort, 1992; Scully, 1989; Vrooman, 1995; Zimbalist, 1992a) have detected no impact of free

[21] Krautmann argues that the stochastic nature of productivity is the source of apparent contract disincentive effects and that 'these vague allegations of shirking are merely a statistical artifact' (1990, p. 967).

agency on competitive balance.[22] The contradiction can be resolved with the realization that, while player movement and competitive imbalance are connected in theory, talent acquisitions do not necessarily imply dominance if those acquisitions are inefficient. Free agent acquisitions are inefficient if the irreversibility proposition holds, or if free agents are paid more than they are worth. To further complicate matters, there are three different issues in the conceptualization of competitive balance: the dominance of large-market clubs, the closeness of league competition within the season, and the continuity of performance (superior or inferior) from season to season. The purpose of this section is to untangle and reconcile empirical evidence on the effects of free agency on competitive balance.

A virtual league. The competitive balance of hypothetical leagues consisting of average large and small market teams can be simulated from estimates of revenue and cost functions similar to those specified in the general theory above. Unfortunately, revenue and cost data released by team owners in North American sports leagues are obscure and unreliable. Consistent and reliable data are available for estimates of revenue and cost equations from a series of annual reports by *Financial World*, now published by *Forbes*. Parameters of the revenue and cost equations implicit in the general model were estimated (Vrooman, 1995) using the double-log specification:

$$\ln R_{it} = \beta_0 + \beta_1 \ln p_{it} + \beta_2 \ln w_{it} + \beta_3 t_{it} + e_{it}$$

$$\ln C_{it} = \gamma_0 + \gamma_1 \ln p_{it} + \gamma_2 \ln w_{it} + \gamma_3 t_{it} + e_{it}$$

where β_3 and γ_3 are the exogenous rates of growth in revenue and costs for team i over the 3 years of the survey 1990–1992, p_{it} is the 1990 population of the home market SMSA and w_{it} is the average winning percentage for the current and two previous seasons.[23] Based on estimates of this model, a large market revenue advantage is the greatest in MLB ($\beta_1 = \cdot 28$), less for the NBA ($\beta_1 = \cdot 18$), and almost nonexistent for the NFL ($\beta_1 = \cdot 03$). MLB has the highest revenue winning elasticity ($\beta_2 = \cdot 60$), the NBA has slightly less ($\beta_2 = \cdot 49$), while the NFL

[22] QE (1974) find no relationship between MLB home market population rank and win percentage rank in the period 1900–70. Scully (1989) finds no significant relationship between population and winning percentage in either the pre-free agency (1961–76) or post-free agency periods (1977–87); but finds a significant drop in the average standard deviation of winning percentages for the NL between pre- and post-free agency periods. Zimbalist concludes, 'By any measure competitive balance has not only not become more unequal since 1976, it has become noticeably more equal' (1992a, p. 95). Based on four measures of competitive balance (including dispersion of wins), QF conclude that pre- and post-free agency periods are indistinguishable and that 'the experience in baseball since 1976 looks almost exactly like the prediction of microeconomic theory, that is no change in competitive balance due to free agency' (1992, p. 284).

[23] 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 settlement for players in 1990. Because of 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.

has the lowest revenue elasticity of winning ($\beta_2 = \cdot 12$).[24] Estimates of the cost functions indicate a higher cost per unit of talent in the large markets in MLB $\gamma_1 = \cdot 22$,[25] and as expected, the winning elasticity of cost is significantly greater than unity ($\gamma_2 = 1 \cdot 30$) only for MLB. This implies that the increasing marginal cost (diminishing marginal returns to talent) conditions required for profit maximization exist only in MLB, and supports the contention that the large market teams are paying significantly more for the same level of talent. The naive revenue model alone should yield accurate predictions for the NBA and the NFL at C in Figure 1, because of the severe restrictions on costs in the NBA and NFL. Free agency in MLB introduces the marginal cost of talent (γ_2) and diseconomies of market size (γ_1) into the profit maximization calculus and yields a competitive solution at $A_1 - A_2$ that is clearly more balanced than revenue maximization at C.[26]

A random event. The actual competitive balance within real leagues was comparatively analyzed in Vrooman (1995, p. 984) using a normalized standard deviation index suggested by Scully (1989), Noll (1991), and QF (1974).[27] Although there are several league-specific factors inherent in the respective games that might affect competitive balance, some relevant intraleague and interleague observations can be made.[28] First, rather than a decreased competitive balance in MLB, as suspected by free-agency opponents, both MLB leagues became increasingly competitive during the pre-strike free-agency period (1976–1992). Second, the NBA has been the least balanced of all three leagues, and there was no tendency for balance to be increased by the imposition of the salary cap in 1984. Third, the NFL has been the most balanced of the three leagues, and there has been little change in that balance over the period studied. These results are consistent with the findings of Scully (1989), Noll (1991) and QF (1992). The NFL has been the most competitively balanced, MLB has been less balanced, and the

[24] All of these estimates are statistically significant at ·01. Based on the 1970 and 1971 seasons, Noll (1974) concluded that market size was most important in MLB, less important in the NFL and least important in the NBA. Winning, as a determinant of home attendance, was 'extremely important' in the NBA ($\beta = \cdot 8$), moderately important in the NFL ($\beta = \cdot 6$), and least important in MLB ($\beta = \cdot 4$).

[25] This implies that $MC_1 > MC_2$ at ·500 in Figure 1.

[26] Substitution of the estimates of β_1 and β_2, and the average populations for small and large markets into the revenue maximization solution 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. An unbiased competitive balance solution of the general model for MLB using cost parameters yields the corrected large/small competitive balance of ·517/·483 for MLB.

[27] The measure of competitive balance normalizes the *actual* standard deviation of winning percentages in the league with an *idealized* standard deviation that 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 an n game season of $\cdot 5/\sqrt{n}$. If the ratio of actual to idealized is one, then the league is competitively balanced.

[28] Interleague comparisons of win variance are problematic because of injury or performance risk differences in the games. For example, there are twelve players on the active NBA team roster, twenty-five for MLB and forty-five for the NFL. Competition variances may reflect the greater importance of a few players for team production on smaller rosters. This risk hypothesis would be consistent with low win variance in the NFL and high variance in the NBA.

NBA has been the least competitively balanced. This is consistent with the virtual-league simulations of the theory developed above.

Player movement. According to the theory presented in Figure 1, free agent movement is governed by two counteractive forces. First, players are expected, *ceteris paribus*, to move from winning clubs to losing clubs, because the contribution of an individual player is necessarily greater to the marginal product of a losing team than it is to a winning team. Abstracting from market size, all winning teams have a tendency to lose talent to those clubs that they dominate. For winning clubs, interseasonal talent redistribution is consistent with the move toward B in Figure 1. This talent flow is dampened, however, by revenue imbalances among teams, and free-agent talent is drawn to large market dominance at $A_1 - A_2$. A small market club is expected to suffer net losses of free agent talent to a large market club, regardless of team quality. Over the course of interseasonal talent redistribution, the most talented large market clubs may lose free agents but they replace them, and the least talented small market teams may reacquire talent, but not enough to improve their prospects of winning. The heaviest net losses occur for small market winning teams and the heaviest net gains, for large market losing teams. In theory, the cold truth of MLB competition should lie in the large market dominance at $A_1 - A_2$.

Regression estimates of the factors affecting free agent losses and gains in the prestrike/postcollusion interim 1990–93 are reported in Vrooman (1996, p. 985). Free-agent losses, gains and net gains are each specified as a function of the club's winning percentage in the preceding season, with binary variables for the eight smallest revenue and eight largest revenue franchises.[29] The results are consistent with the predictions of theory and the findings of previous empirical studies.[30] First, the loss of free-agents is directly related to the overall quality of a team's talent, as measured by its winning percentage in the preceding season. This is especially true for winning small-market teams. Second, the acquisition of free-agents is inversely related to team quality, and there is strong evidence that large-revenue clubs, regardless if they are winning, acquire

[29] Free agent gains and losses are the annualized new salaries (with bonuses) of free agents who switched teams during 1990–93 and remained on the opening rosters of their new teams for the next season. Net free-agent gains are the differences between the new salaries of free agents gained and of free agents lost. Due to the growth in media and stadium revenues, the usual distinction between large and small based on population is misleading. Five of the large revenue clubs have a media advantage (Yankees, Mets, Dodgers, Red Sox and Braves) and three played in new stadiums during the period studied. The Toronto Blue Jays played their first full season in the Sky Dome in 1990; the Chicago White Sox first season in the new Comiskey was 1991, and the Baltimore Orioles opened at Camden Yard in 1992.

[30] Cymrot (1983) used Scully-like MRPs as dependent variables rather than player salaries and found that 1977–79 'nonmarginal' free agents tended to move from good teams in small markets to poor teams in large markets. This led to the conclusion that size by itself does not imply domination due to the tendency for high quality teams to lose free agents. Daly and Moore concluded that 'the effective termination of the reserve clause for veteran players in 1976 has been followed by a series of free agent transactions which have clearly strengthened big city teams' (1981, p. 94). Quirk and El Hodiri detect 'a strong positive correlation between the strength of an AL team (1929–50) and its payment for players ... The NL data are particularly striking: if Rickey's teams, the Dodgers and the Cardinals, ... are excluded, a strong positive correlation between acquisition costs and percentage of games won emerges' (1974, p. 52).

significantly more free agent salaries than the rest of the league. Small-revenue clubs lose free agents that they cannot replace. Large-market clubs, regardless of team quality, acquire significantly more net free agents than the rest of the league, while winning teams, especially those in small-revenue markets, suffer the most significant free-agent losses. It is important to realize that movement of free agent *salaries* among clubs does not necessarily imply a corresponding redistribution of *talent* if players are not paid the value of their marginal revenue product. The question then arises as to whether this observed movement of free-agent *salaries* actually reflects a redistribution of talent.

A random walk. As it relates to the effects of interseasonal free-agent talent acquisition, the most important aspect of competitive balance concerns season-to-season performance continuity of individual clubs. If the league is competitively balanced in this respect, then dynasties and doormats are the exception, and team fortunes are random walks from season to season. Analysis of the season-to-season continuity of team performance during and immediately preceding the free-agency era involves the estimation of an autoregressive model for winning percentage: $w_{it} = \beta_0 + \beta_1 w_{it-1} + e_{it}$ for the ith team in season t. If the outcome of a league is predictable, then $\beta_0 = 0$ and $\beta_1 = 1$ and if the league is a random walk, $\beta_0 = \cdot 5$ and $\beta_1 = 0$. Estimates of the model are presented in Vrooman (1996, p. 352) for pre-free agency (1970–76) and free agency periods (1977–1993).[31] The results are consistent with the consensus of other studies that competitive balance has significantly increased in MLB during free agency.[32] Immediately preceding free agency, roughly 70% of a club's performance is 'predetermined' by its record in the preceding season ($\beta_0 = \cdot 142, \beta_1 = \cdot 715$). Estimates for the succeeding periods reveal a gradual erosion of season-to-season continuity of performance and a significant increase in competitive balance from 1977 to 1989. The most striking result, however, is significant evidence of absolute discontinuity of performance of MLB teams ($\beta_0 = \cdot 480, \beta_1 = \cdot 036$) during the pre-strike/post-collusion interim 1990–1993. In the face of renewed free-agent acquisitions, MLB had become a season-to-season random walk.[33]

If higher-salaried free agents continue to move to large markets, then the revenue advantage seems to remain a significant factor in the acquisition of player talent. But if observed free agent talent redistribution fails to result in an increased dominance of the clubs that acquire talent, then the acquisition of talent is *necessarily* inefficient. This implies that free agents are being paid more than they are worth. Consequently, the acquisition of free-agent talent in a monopolistic environment could significantly increase a large market club's

[31] The free agency era is punctuated by two watershed events (1981 strike and 1986–89 collusion) that subdivide it into post-free agency (1978–80), pre-collusion (1983–85), collusion (1986–89) and post-collusion (1990–93) periods.

[32] Based on a comparison of 1962–76 and 1977–89 periods, Balfour and Porter conclude that 'a team's performance is *less* likely to be repeated in subsequent years if players are free to move among teams' (1991, p. 17).

[33] Estimates by students in a sports economics seminar at Rice University indicate that MLB continuity (imbalance) has been restored after the 1994 strike. Estimates by students in a sports economics seminar at Vanderbilt University show a sudden discontinuity of performance in the NFL for the 1999 season.

marginal cost of talent to the point where its revenue advantages are negated.[34] In this case, the random walk is the result of winning small-market teams being dismantled and not reassembled in their large-market destinations.

Pay and performance. Following Scully's methodology (1974), several studies have tried to measure the degree of monopsonistic exploitation of MLB players by comparisons of observed player salaries with various individual performance measures.[35] In Scully's attempts to find the worth of a player before free agency (1968–69), he concluded that:

> the economic loss to professional ball players under the reserve clause is of a considerable magnitude (1974, pp. 929).

In the early years of free agency, several studies replicated Scully's results through comparisons of the rates of exploitation (ratios of player salary to marginal revenue product MRP) between free agents and non-free agents. While there is significant evidence of underpayment of non free agents in each of these studies, there is also evidence that free agents were consistently overpaid relative to their MRPs.[36]

Bruggink and Rose (1990) use Scully's method to find the amount that MLB free-agent salaries were reduced by the collusion of MLB owners following the 1985 and 1986 seasons. While they reach the conclusion that the salary/MRP exploitation ratios were 28–38% lower during the period of alleged collusion, they also find that pre-collusion free agents in 1984 were paid 22% more than their MRP's (Bruggink, 1990, p. 1037). Blass (1992) estimates that 76 to 86% of 1985–6 position players (non-pitchers) eligible for free agency are overpaid. Zimbalist (1992a, 1992b) observes that, while players with less than 6 years of experience are paid between one-fifth and two-thirds of their MRP's, players with six or more years of MLB experience were paid on average 23·6% above their MRP in 1986, 31·5% above their MRP in 1987, 28·0% above in 1988, and 39·7% above in 1989.

[34] Zimbalist (1992), observes a reduction in the coefficient of variation (μ/σ) among MLB team average player salaries from the free agency period through the post-collusion period: $\mu/\sigma = \cdot 357$ in 1978 to ·294 in 1986 and ·225 in 1990. In the monopolistic environment of the post-collusion period, however, the trend is reversed: $\mu/\sigma = \cdot 304$ in 1991, ·306 in 1992, and ·333 in 1993. Team player costs were becoming more widely dispersed than team revenues. Over the four-year period 1990–93 the 86% explosion in team average salary from \$597 K to \$1·109 M was more than doubled by the 175% increase in the standard deviation of team average salary from \$134 K to \$369 K.

[35] Scully's method characterizes tautological human capital specifications of marginal productivity theory. The methodology explicitly assumes that 'individual performance carries with it no externalities, so that team performance is simply the linear summation of individual performance' (1974, p. 921). Later, Porter and Scully (1990) and Scully (1989, p. 36) suggest that residual productivity should be considered a measure of a team manager's marginal revenue product.

[36] Sommers and Quinton find evidence (1982, p. 434, Table II) that free agents are worth more in large markets, but struggle with the conclusion that 'These estimates of net benefits and costs suggest, rather misleadingly, that virtually all of the free agents in 1977 were grossly overpaid ... it seems that owners could get better buys on talent by ignoring free agents' (1982, pp. 433–34). Raimondo finds evidence (1983, p. 185, Table I) of exploitation for mediocre 1977 free agents but that star players and average free agent pitchers received more than their MRP's.

Zimbalist concludes:

> Team owners since 1976 have done a singularly unimpressive job of signing top-performing free agents or paying a player according to his output. Consequently, average team salary has been related only tenuously to team performance (1992a, p. 91).

The implications of each of these studies lead to the conclusion that limited free agent players are systematically paid more than they are worth, not because of their freedom, but because of their artificial monopoly power.

Policy conclusions for MLB

The free agency fix. As the baseball players' labour market has evolved from monopsony into bilateral monopoly during the free agency era, it has been segmented into three seniority tiers by different combinations of monopsonistic and monopolistic forces. The lowest tier of players, who are not eligible for either free agency or salary arbitration, are controlled by the vestigial monopsonistic power of the club owners, whereas the upper tier is dominated by the artificial monopoly power of players eligible for limited free agency. Players in the middle tier are subject to the binding constraints of the reserve clause, but are eligible for salary arbitration.[37] Over the free agency period, the relative distribution of players among the tiers has remained stable. Two-thirds of MLB players are confined to tiers one and two (about one-fourth are eligible for arbitration in tier two) and one-third of the players are eligible for free agency. In the period immediately before the 1994 strike, player salaries were increasing at 5–10% in the lower tier, compared to an annual growth rate of about 25% for all players. Any proposed 'free-agency fix' for this labour market twist should involve a reduction in the eligibility criterion for free agency in combination with the elimination of the salary arbitration procedure that defines tier two. Elimination of tier two simultaneously reduces the domain of monopsonistic exploitation and weakens the monopoly power held by free-agent eligible players.[38] The

[37] Since 1991, players with 3 or more years of MLB experience and the top 17% of the two-year service class are eligible for arbitration. Salary arbitration eligibility had been 2 years from 1974–85 and 3 years from 1986–90.

[38] It also removes an arbitrary procedure that worsens competitive balance within the league. The stated purpose of arbitration is to achieve salary parity among players of equal talent and experience regardless of market size. MLB arbitration criteria may include performance measures and salary information about other players in the league, but not information concerning the financial position of the player and the club. Small market teams are forced to pay players partially what they would receive in the large revenue markets, and large market clubs are allowed to pay players partially what they would receive in small markets. The parity criteria of the MLB arbitration procedure may reduce the monopsonistic exploitation by small market owners, but it also reinforces the competitive advantage for large market clubs in the second tier Scully (1989) argues that relatively weaker franchises are more likely to go to arbitration, but makes the inference that they are the major beneficiaries of the process. This is not necessarily true. More than 90% of arbitration proceedings are initiated by players. Small market clubs are involved in salary arbitration because they are more likely to be *brought* to arbitration by small market players seeking parity.

major empirical question concerns the optimal experience requirement for a MLB player to assume free-agent eligibility. If the conditions for free agency were unlimited, there would be no incentive for clubs to internalize the development of the skills of young players. To be effective, the free agency fix should reduce the free-agent eligibility requirement precisely to the length of MLB service necessary for the MLB clubs to recoup their investment in player development costs.

Player development expenditures (PDX). Although detailed expenditures are sporadically reported for the league or the individual clubs, the average PDX per club in 1994 can be estimated. An exponential growth rate of 12·6% for PDX in the free agency period (1977–91) was used to project a per team average PDX of \$9·200 million in 1994.[39] The relative efficiency of the overall farm system was then gauged by a comparison of PDX with the number of rookies placed on the 25-man active roster of the parent MLB club. Of the 746 active major leaguers opening the 1994 season, 96 were rookies. This yields an average of 3·43 rookies for the 28 clubs and a discounted PDX of \$2·53 million per rookie produced by the MLB player development system.[40]

Development cost recovery (DCR). The monopsony conditions of tiers one and two provide MLB clubs with an exploitation surplus more than sufficient to recover PDX.[41] The free agency eligibility criterion should be set where the present value of the exploitation stream under modified reserve is equivalent to the discounted PDX of \$2·53 million in 1994. The exploitation surplus for MLB players in 1994 was isolated in (Vrooman, 1995) with the estimation of an exponential salary model with performance measures omitted:

$$\begin{array}{lllll} \ln S = \cdot 666 - 1{\cdot}269\, T_1 - \cdot 190\, T_2 + \cdot 240\, X - \cdot 012\, X^2 & & & & R^2 = \cdot 633 \\ \quad (-3{\cdot}73) \quad (-8{\cdot}18) & (-1{\cdot}50) & (6{\cdot}17) & (-6{\cdot}08) & N = 746 \end{array}$$

where S is a MLB player's annual salary, T_1 is a binary variable for players of tier one, T_2 is a binary variable for players in tier two with less than four years

[39] This estimate is consistent with actual PDX reported for the Pittsburgh Pirates of \$9·163 million and Baltimore Orioles of \$8·768 million in fiscal 1993. Data derived from Scully (1989, p. 118) yield a per club average PDX of \$1·029 million in 1977, \$1·546 million in 1980 and \$2·180 million in 1983. Data from Zimbalist (1992, p. 113) yield a club average PDX of \$1·538 million in 1981, \$2·180 million in 1983, \$2·846 million in 1986, and \$5·5 million in 1989. Expenses released by the MLB Commissioner's Office show an average PDX of \$7·200 million per team in 1991.

[40] A nondiscounted average PDX of \$2·68 per player is discounted assuming a growth rate of 12·6%, a discount rate of 10%, and the average rookie spends 4 years in the minors.

[41] Zimbalist (1992a, 1992b) estimates Scully-like equations for each of the three segments of the MLB labour market, which he labels apprentices, journeymen and masters. In his results, Tier 1 players are paid 21·8% of their MRPs in 1986, 17·1% in 1987, 22·6% in 1988 and 24·3% of their MRPs in 1989. By comparison, Tier 2 players receive 50 to 65% of their MRPs and free-agent salaries exceeded MRPs by 23 to 39% over the same period. Based on these estimates Tier 1 players generate an exploitation surplus of 3·59 times their salaries in 1986, 4·85 times their salaries in 1987, 3·42 times their salaries in 1988, and 3·12 times their salaries in 1989.

experience, and X is actual years of MLB service for all players on opening day rosters in 1994.[42] These results suggest that the 1994 vintage of rookies would generate an exploitation surplus of approximately 2·56 times their discounted salaries for the years in which they are subject to the modified reserve clause.[43] Under these conditions, a player under reserve would generate a present value DCR of \$1·7 million (two-thirds of discounted PDX) following his third year, and following his fourth year, he would repay the club all of his player development expense with a present value DCR of \$2·53 million.[44] Full DCR requires that free agent eligibility be set at four complete years of MLB service.[45] Under the proposed free agency fix the proportion of players facing monopsony exploitation would be reduced from two-thirds to one-half, while the number of players in the free agent pool would increase from one-third to one-half of active players.

II THE OWNERS' CAPITAL MARKET

The economics of professional sports has been relegated to the realm of labour theory by the assumption that owners of sports franchises are single-minded maximizers of profit *in vacuo*. According to theory developed within this vacuum (including the general theory above), free agency in a sports league would not affect the *real* decisions of teams, but it would result in a transfer of incomes from owners to players in the form of reduced exploitation and higher player salaries. As predicted, player cost as a share of league revenues has risen from 20% to 58% over the first two decades of free agency in MLB.[46] The escalation in player costs has been attributed, by association, to the institution of free agency. Economic theory has overlooked a myriad of other contributing factors, such as an explosion in media rights fees and an underlying revolution in

[42] T_1 is a binary variable for players with less than 2·13 years MLB service (not eligible for salary arbitration), T_2 is a binary variable for players with at least 2·13 years but less than 4 years of MLB service (third year players and the top 17% of second year players), X is the actual years of MLB service and S is the annual salary for 476 players on active rosters opening the 1994 season ('t' ratios in parentheses). Using a similar model, Blass (1992) shows that wages increase with experience independent of productivity.

[43] Monopsony conditions of tiers 1 and 2 provide MLB clubs with an exploitation surplus of λS for the mean salary S where $\lambda = (1-\alpha)/\alpha$ and $\alpha = S/MRP$. Exponentials of the T_1 and T_2 coefficients are direct estimates of α: $\alpha_1 = \exp(-1{\cdot}269) = {\cdot}281$ for tier 1 and $\alpha_2 = \exp(-{\cdot}190) = {\cdot}827$ for first-year arbitration-eligible players.

[44] If salaries for players under reserve grow at their current annual rate of 8%, and if the exploitation surplus is discounted at a 10% rate, then the present value of a 1994 player's contribution to the surplus can was at \$324·9 thousand per player in year 1, \$521·3 thousand in year 2, \$846·5 thousand in year 3, and \$835·3 thousand in year 4.

[45] The 1990 *Basic Agreement* (*XXIV*) established a 'Baseball Economic Study Committee' that recommended that salary arbitration be dropped in exchange for free agency after three years. Scully (1995, p. 59) estimates that over the free agency period (1977–92) a MLB player's MRP exceeds his earnings after 547 career games (about 3·6 years).

[46] After player costs increased rapidly from 20% of the MLB revenue in 1977 to 40% in 1981, the ratio drifted down to 32% in 1989 due largely to the collusion of MLB owners. Following the inordinately large MLB media rights fees agreement for 1990–93, the players' share of revenue jumped to 58% by the strike of 1994.

ownership structure. During the first 20 years of free agency, over 40 MLB franchise ownerships were coincidentally transferred amid a rapid transition of ownership structure from sole sportsmen to leveraged syndicate.[47] It is entirely possible that salary escalation during the free agency period was driven by a capital-market game of leveraged syndication being actively played out on a passive labour-market field of restricted free agency. The most formidable problem with developing a theory to unify capital and labour markets stems from Modigliani and Miller's *irrelevance proposition* (1958) that the real decisions of a firm (team) are independent of its financial structure. The weakness of the *irrelevance proposition*, however, is that the separation of financial and real decisions rest precariously on the conventional assumption of profit maximization. If team owners are more generally viewed as sportsmen, who are willing to sacrifice profit in pursuit of winning, then it can be shown that the owners' capital-market decisions are inseparably linked to the baseball players' labour market (Vrooman, 1997a).

The irrelevance proposition

A unified theory of capital and labour markets has been deemed difficult, if not impossible, following the Modigliani and Miller's (MM) (1958) famous *separation theorem*, in two parts. The first separation principle holds that:

> *the market value of any firm is independent of its capital structure*... ie, *the average cost of capital to any firm is completely independent of its capital structure and is equal to the capitalization rate of a pure equity stream of its class* (emphasis original) (1958, pp. 268–69).

The second separation principle concludes that optimal operating decisions of the firm do not depend on its financing decisions, and that *operating and financing decisions are separable*. The unified theory of capital and labor markets developed in Vrooman (1997a) relaxes the underlying assumption of the second separation principle: that a MLB franchise is always acting in the best interest of equity and debt holders (Modigliani and Miller, 1958, p. 288). The theory follows a path blazed by Scitovszky (1943) and later redeveloped through the agency theory of Jensen and Meckling (1976).

According to the first separation principle, leverage is irrelevant to the valuation of the franchise, because any increase in the value of the club due to the use of cheaper risk-free debt is exactly offset (according to the law of conservation of risk) by an increase in the cost of equity. If each sports franchise is subject to the same business risk, then all combinations of leverage and syndication are equivalent to the capitalization of 'a pure equity stream of its (risk) class'. The implications of the first separation principle can be shown

[47] Media revenues increased from an average of $2 million to $26 million per club during the first two decades of free agency. Over the same period, franchise prices and player salaries have both risen by 15%. In 1976 Seattle and Toronto paid franchise fees of $7 million, and in 1995 Arizona and Tampa paid an $130 million expansion fees.

through simple valuation models of sports clubs of the risk class ε. The value of a pure-equity team is simply π/r, where net cash flow π is capitalized by $r = r^* + \varepsilon$, a risk-free rate r^* adjusted for league risk ε shared by all clubs. The value of a leveraged franchise is the sum of outstanding debt D and equity X shares:

$$V = D + X = D + (\pi - r^* D)/\rho$$

where $X = \pi - r^* D$ is net cash flow after interest capitalized at a rate ρ reflecting the financial risk of leverage. If $D = \lambda V$ and $\rho = r^* + \varepsilon/(1 - \lambda)$ is r^* adjusted for risk ε concentrated on the unlevered portion of the franchise $(1 - \lambda)V$, then the value of a levered franchise becomes the expected value of 'the pure equity stream of its class':

$$V = D + (\pi - r^* D)/\rho = \pi/r.$$

According to the second separation principle, the operation of the team is independent of its financial or ownership structure, *but only if* each owner maximizes the value of his franchise with respect to winning. If team owners are more generally perceived as 'sportsmen', who jointly maximize franchise values and the satisfaction derived from winning, then the second separation principle does not necessarily follow from the first. The *irrelevance proposition* obtains only in the exceptional case of the pure profit maximizer. In the more general optimization of the multiple objectives of the sportsman owner, the capital and labour market choices of team owners become linked in a manner that influences the on-field performance of their teams.

The sportsman effect

The optimization problem facing the sportsman owner concerns the joint maximization of franchise value (V) and the satisfaction (S) derived from winning (w). The sportsman reaches an optimum where the marginal value of the franchise with respect to winning is equal to the negative of the rate of substitution (RS) between winning and franchise value: $MV = (MR - MC)/r = -RS$. Compare this general sportsman equilibrium at B in Figure 2 with the customary, but limiting case of pure profit maximization (for $RS = 0$) at A. This leads to the proposition that a sportsman owner sacrifices franchise value for winning and expands the talent of his club beyond its profit maximum. The resulting undervaluation of the franchise measures the *sportsman effect*.

The sportsman optimum at B would also maximize the residual of franchise value above a minimum value necessary to keep the owner in the league. The indifference curve S_0 in Figure 2 represents the minimum satisfaction that will keep the sportsman in the game. The vertical distance beneath S_0 is the franchise value necessary to compensate the owner for the on field performance of the team, and the difference between franchise value V_0 and minimum satisfaction S_0 (the envelope $\omega - \omega'$) is the residual value that is maximized by the

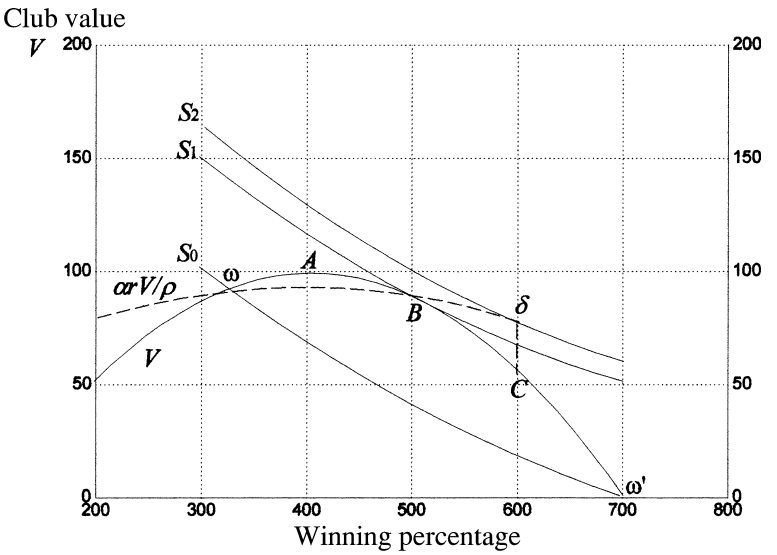

Figure 2. Sportsman and Steinbrenner effects.

sportsman.[48] For winning percentages above ω, winning no longer compensates the owner for the loss in franchise value, and for winning percentages below ω', franchise value no longer compensates the sportsman for losing. If the owner's satisfaction were to fall below S_0, the sportsman would sell the sports franchise.

As was the case in the general theory, a unique equilibrium in the unified theory is insured by the zero-sum constraints of league competition ($\Sigma w_i = n/2$), and the equilibrating mechanism for the simultaneous solution of a league of sportsmen is the cost per unit of talent faced by all teams in the league. If $\Sigma w_i > n/2$, then the excess demand for talent within the league will force the cost per unit of talent to increase and the cash flow and values of all clubs to decrease. Conversely, if $\Sigma w_i < n/2$, then the excess supply of talent will depress player costs and increase cash flows and values of all clubs until a league equilibrium is reached consistent with $\Sigma w_i = n/2$. League equilibrium for a simplified two-team league requires the simultaneous optimization of both clubs $(MR_1 - MC_1)/r + RS_1 = (MR_2 - MC_2)/r + RS_2 = 0$. This is shown in Figure 2 at B for one of two cloned sportsmen, who compete in a simplified league of equal revenue ($MR_1 = MR_2$ at ·500) clubs.

Several distinctions between a sportsman league ($\partial S/\partial w > 0$) and a pure-profit league ($\partial S/\partial w = 0$) can be drawn by constraining $RS_1 = RS_2 = 0$ for the limiting case of a profit maximizer. If owners are seen equal sportsmen ($RS_1 = RS_2$ at ·500), then average player costs in a sportsmen league are rRS

[48] In a classic argument, Scitovszky (1943) noted that the maximum residual would occur at the same winning percentage as the satisfaction optimum B if the slopes of the indifference curves are the same at that percentage In order for S_0 to be a vertical displacement of S_1, the sportsman effect must be independent of the value of the franchise.

higher at ·500 than for clubs in pure-profit league ($RS = 0$). As a result, the value of a franchise in a sportsman league would be RS at ·500 lower than the same club in a pure-profit league. This leads to the conclusion that, in a sportsman league, player costs will be higher for all teams, and all franchises will be undervalued, compared to the same clubs in a profit maximization league. The resulting franchise undervaluation (vertical difference between A and B in Figure 2) serves as a measure of the sportsman effect in league equilibrium.[49]

If owners of unequal revenue clubs ($MR_1 > MR_2$ at ·500) are identical sportsmen and the rates of substitution between winning and franchise values for both are symmetrical around ·500, then the large revenue club encounters rates of substitution above ·500 that are relatively smaller ($RS_1 < RS_2$) than those for the small revenue club. If the rate of substitution is greater for the small revenue club than for the large revenue club, then the advantage of a large revenue franchises is mitigated. The relatively smaller rates of substitution between winning and club value for large-revenue sportsmen imply that a sportsman league of unequal revenue clubs is more competitively balanced than is a profit maximization league comprised of the same clubs, and that the degree of competitive balance is proportional to the relative incidence of sportsmen owners in the league.[50]

The Steinbrenner effect

The joint objectives of winning and franchise value for the sportsman foster a unique environment in which separation of ownership and control of a franchise can create potential agency effects.[51] A sportsman owner, who syndicates $(1-\alpha)$ limited partnership LP shares and retains an α managing general partnership GP share, can gain all the satisfaction from (credit for) winning, while bearing only an α fraction of the reduction in franchise value. The optimization solution $\alpha(MR_i - MC_i) = -rRS_i$ for a GP sportsman implies that the cost of winning in terms of reduction of franchise value is diluted by ownership syndication, and that the GP can improve his position at the partial expense of the LP shares. Compare the sole-sportsman optimum at B to the unlevered ($\rho = r$), syndicated ($\alpha < 1$) GP-sportsman optimum at δ in Figure 2. Maximization of the

[49] Partial sportsmen ownership within the league will increase player costs for all clubs, regardless of different ownership objectives, because league clubs share the same talent pool. This leads to the conclusion that partial composition of league ownership by sportsmen will result in a proportional increase in average player costs for all clubs in the league, and the values of all clubs will be reduced in direct proportion to the concentration of sportsmen owners.

[50] Sportsman league equilibrium yields $(MR_1 - MC_1) - (MR_2 - MC_2) = r(RS_2 - RS_1)$, rather than a greater dominance of large-revenue teams implied by $MR_1 - MC_1 - (MR_2 - MC_2) = 0$ under conditions of pure profit maximization. The greater the divergence between MR_1 and MR_2 at league equilibrium, the less the competitive imbalance. If the marginal product of talent is symmetrical between teams, then $MC_1 = MC_2$ at ·500 and $MC_1 > MC_2$ above ·500. This implies $MR_1 > MR_2$ at league equilibrium which is more competitively balanced than $MR_1 = MR_2$ solutions.

[51] Agency effects do not exist under the special conditions of profit maximization. An owner's maximization of his own equity share $\alpha \partial V/\partial w = 0$ is equivalent to the value maximum $\partial V/\partial w = 0$ in Figure 2.

satisfaction derived from the GP's share (or maximization of the residual envelope $\alpha V - S_0$) at δ drives the cost of winning up and the value of the franchise down to C for all partners, while allowing the GP to reach a higher level of satisfaction along S_2. Hence, the *Steinbrenner effect* implies that a syndicated sportsman owner has an incentive to expand the talent level of a club beyond that of the sole-sportsman owner.[52] If the syndicated sportsman disturbs a league previously in equilibrium, then $\Sigma w_\iota > n/2$ and ownership syndication would create an excess demand for talent within the league. The average cost of talent would rise and franchise values would fall for all clubs until $\Sigma w_i = n/2$. (Marginal cost would increase and the value function would shift downward and to the left in Figure 2). All clubs must now pay higher player costs in direct proportion to the number of Steinbrenner clones in the league. In a league of teams owned by syndicated sportsmen, player costs will be higher for all teams and all franchises will be undervalued compared to the same clubs in either a pure-profit or sole-sportsman league. This undervaluation measures the full *Steinbrenner effect* in league equilibrium.[53]

Other peoples money

In addition to syndicating ownership, a sportsman owner can also choose any combination of financial leverage. A buyer may choose to acquire a franchise by borrowing a portion of the purchase price or by assuming outstanding prior debt as part of the purchase price. A syndicated sportsman GP may decide to buy out limited partners by leveraging the franchise. In general, team owners with varying degrees of leverage and syndication would jointly maximize their preference for winning and the value of their equity share in the franchise such that: $\alpha r/\rho MV + RS = 0$, where $\rho = r^* + \varepsilon/(1-\lambda)$ and $r = r^* + \varepsilon$. In the case of the levered ($\rho > r$), sole-sportsman ($\alpha = 1$), maximization of the equity share yields $r/\rho MV + RS = 0$. The cost of winning in terms of lost potential equity is reduced by $(1 - r/\rho)$ due to the higher financial risk ρ concentrated on the levered equity share. In Figure 2, a non-syndicated levered sportsman is constrained by $r/\rho MV$, rather than MV for the unlevered sole sportsman, and the levered sportsman can reach a higher level of satisfaction at δ along S_2, rather than S_0. The levered sportsman has *Steinbrenner*-like incentives to acquire

[52] The New York Yankees were purchased from CBS for $10 million in 1973 by a syndicate headed by the general partner, George Steinbrenner, who was also chairman of American Shipbuilding. Following the purchase the new GP assured that 'We plan absentee ownership as far as running the Yankees is concerned. We're not going to pretend we are something we aren't. I'll stick to building ships'. Since that time, however, Steinbrenner obtained majority 55% ownership in 1976, and 'the Boss' has become the most visible and intrusive owner in MLB. A rational investor would obviously value the franchise at C. The Yankees, for example, have had interested outright buyers, while the original LP's have had difficulty finding a market for their limited shares with Steinbrenner as GP.

[53] It can be shown that a syndicated sportsman league of unequal revenue clubs is more competitively balanced than is either a pure-profit maximization league or a sole sportsman league comprised of the same clubs. It is also true that a league of syndicated franchises will face a closing window of operation compared to the range of existence for the same clubs in either a pure-profit or sole-sportsman league environment (Vrooman, 1997a).

more talent and team wins than does his unlevered adversary with an equal revenue base and ownership structure.[54]

Although leverage may initially lead to *Steinbrenner* effects, the operation of the leveraged franchise is different from the syndicated club in at least one positive respect. According to the 'control hypothesis' of debt creation, the threat caused by the failure to make debt service payments serves as an effective motivating force to make organizations more efficient. For example, the levered sportsman in Figure 2 would operate at δ only if the amount of debt (D) in the franchise was less than C. For debt levels above C, the *Steinbrenner effect* at δ will be controlled by the legal obligations of franchise solvency $V = D$. Under syndication, the GP sportsman may expropriate value from his limited partners, but in highly leveraged financial structures, the levered sportsman is restricted. The solvency requirement $V = D$ above C, therefore, dominates the Steinbrenner optimum at δ. At successively higher levels of debt above C, a highly leveraged club's talent level is trimmed, and the *Steinbrenner* effect is reduced as D approaches B. The highly leveraged owner is forced to a lower level of satisfaction along S_1 at B, where the *Steinbrenner effect* is completely negated by the control function of debt.

At relatively low levels of financial leverage, there exists an incentive for a levered sportsman owner to expand the talent level (winning percentage) of his club beyond that of the owner of an unlevered franchise with the same revenue potential. A sufficiently high degree of leverage also exists, however, where the *Steinbrenner effect* is self-constrained by fixed debt requirements, so that the operational decisions of a levered sportsman approach those of an unlevered sportsman. The agency effects of leverage will create an excess demand for playing talent and league disequilibrium $\Sigma w_i > n/2$. Player costs will rise and franchise values will fall for all clubs until $\Sigma w_i = n/2$. It is shown in Vrooman (1997a) that sports leagues in equilibrium will have the highest player costs, the greatest competitive balance, and the lowest overall franchise values ranging from pure-profit, to pure sportsmen, to leveraged sportsmen and ultimately to syndicated sportsmen of the same degree.

The compression of franchise values and closure of Scitovszky franchise survival ranges ($\omega - \omega'$) make certain ownership or capital structures vulnerable to the agency effects of increased syndication and/or leverage. The question arises as to which of these financial/ownership structures would prevail over time in a mixed league environment. The possible extinction of different forms of ownership derives from the fact that club owners share a common talent pool.

[54] The levered owner has an added incentive to increase the business risk of the club's operational decisions. Because of the asymmetry of risk between creditors and owners, a leveraged owner will have a strong incentive to acquire players that promise very high payoffs if successful, even if they have a low probability of success. If they turn out well, he captures most of the gains, if they turn out badly the creditors bear most of the costs. For example, the levered owner has an incentive to quick-fix his team through the acquisition of long-shot free agent talent, rather than developing talent through his club's farm system. The levered owner is more likely to engage on the risky side of multiple player trades (several minor leaguers for one veteran), whereas the more risk-averse, unlevered owner would operate on the risk diversification (one for many) side of such trades.

Not only does a syndicated GP expropriate franchise value from his own LP's, but he also shifts part of the cost of the Steinbrenner effect to all clubs in the league. Because of the absence of a debt control in syndicated ownership, the Steinbrenner effect can dominate other financial forms and, at the limit, a league with a large concentration of syndicated Steinbrenner clones will drive other ownership forms from the league.[55] There exists a threshold concentration of syndicated owners in a sports league that will inflate player costs and compress franchise equity to the extent that it will either drive pure sportsman and leveraged sportsman owners from the league, cause them to explore alternative revenues through franchise relocation and stadium extortion or otherwise force them into ownership syndication.

Empirical evidence for MLB

Extinction of the sportsman. Initial support for the theory can be drawn from a brief review of the MLB ownership histories summarized in Vrooman (1997). The MLB free agency period began with an abundance of sportsman owners in league with a few limited partnerships. During the first five years of free agency (1976–1981), player costs as a share of revenues doubled from 20% to 40% and sole sportsmen began to sell their clubs. A transition of MLB from a league of pure to leveraged sportsmen was symbolically completed in 1984, when Calvin Griffith, 'the last of the dinosaurs', sold the Minnesota Twins/Washington Senators (family owned since 1919) for $34 million to financier Carl Pohlad. In the same year, thirty-year owner John Fetzer sold the Detroit Tigers for $53 million to the highly leveraged owner of Domino's Pizza, Tom Monaghan.

The extinction of the sportsman and the proliferation of highly leveraged MLB ownership syndicates during the 1980s were symptomatic of a deeper American leveraged-buy-out (LBO) revolution.[56] In 1989, three ownership transfers were particularly representative of the prevailing LBO influence. First, Eli Jacobs bought the Baltimore Orioles from the estate of Edward Bennett Williams for a bargain price of $70 million. Jacobs had previously executed a series of LBO's in the late 1980s backed by Drexel Burnham Lambert junk

[55] If the value of a franchise were to fall below S_0 and beyond the limits of the Scitovszky existence envelope ω–ω', then ownership of the franchise would be sold. The individual equilibrium of the GP owner at γ creates a league disequilibrium and an excess demand for playing talent. Inflation in player costs would ensue and continue until a syndicated sportsmen league equilibrium was reached for C at ·500. In direct proportion to the concentration of syndicates in the league, the value functions of all clubs in the league would shift downward to the left. Increased player costs and reduced cash flow below S_0 would no longer compensate the sole sportsman for losing, and he would be forced to sell his club outright or to syndicate his current ownership equity.

[56] Since the beginning of free agency, MLB owners had self-imposed an informal '60/40 rule' that sought to restrict the use of leverage in the purchase of a franchise to 40% of its value. After 1983, the rule was to apply to the operation of all clubs at all times. On closer examination, however, the design of the rule was such that the permissible level of debt was actually 40% of the value of the average MLB franchise. According to this '60/40 rule', a small revenue franchise that was valued at two-thirds of the league average could actually maintain a 40/60 equity-debt ratio, whereas a large revenue franchise valued at twice the league average would be restricted to a 80/20 equity-debt ratio.

bonds.[57] Failures of these highly leveraged transactions ultimately forced the sale of the highly profitable Orioles in bankruptcy auction in 1993 for $173 million. Second, Jeff Smulyan purchased the Seattle Mariners for $76·1 million in syndication with Emmis Broadcasting and Morgan Stanley & Company.[58] During the 1980s, Smulyan acquired undervalued radio stations in a series of LBO's through Emmis. The Emmis bottom line fell from a $25 million profit in 1989 to a $23 million loss in 1991, and Smulyan was forced to sell the Mariners in 1992 for the invested capital of $106 million. Third, George W. Bush, in syndication with Rusty Rose and others, bought 58% of the Texas Rangers for $46 million from the highly leveraged oilman Eddie Chiles.[59] Rose and the other investors were heavily involved with Drexel Burnham Lambert junk-bond LBO's of the 1980s.[60]

Although the financial terms of franchise ownership are fully disclosed only on rare occasions, the trend in the 1990s seemed to be toward increased 'leverage' through syndication. For example, in a highly syndicated deal for ownership of the Pittsburgh Pirates, Kevin McClatchy put down only $10 million of the $60 million in required equity.[61] Bush and Rose were the GP's for the group of 15 partners to purchase the Texas Rangers in 1989 for an implicit price of $80 million. Although Bush was the most visible GP until he was elected Governor of Texas, his original investment was only $500 thousand. Similarly, the ownership syndicate that was awarded the Arizona Diamondbacks expansion franchise in 1995 paid an expansion fee of $130 million. Jerry Colangelo ('owner' of the Phoenix Suns NBA franchise), Eddie Lynch and John Teets are the visible GP's, but they invested only $1·1 million in the 20 member investment syndicate. According to the unified theory, the revolution in ownership structure is the protracted consequence of the purchase of the New York Yankees in 1973 by a 15 member syndicate headed by George Steinbrenner.[62]

Sportsman effect revealed. Steinbrenner effects are derived from, and necessarily dependent upon the existence of the sportsman effect. In spite of

[57] Memorex ($523 million) and Telex Corporation ($1 billion) in 1986 and Triangle Pacific ($406 million) in 1988.

[58] Smulyan and Mike Browning each held 10% equity, Emmis owned 47% and Morgan Stanley held 27% in the Mariners and about 25% in Emmis. The amount of debt in the sale was about $35 million.

[59] Chiles bought the Texas Rangers from Brad Corbett in 1980 for $25 million, with $21 million in deferred salaries. Chiles had twice attempted to sell his share to Edward Gaylord who already owned 33% of the Rangers, but was rebuffed by MLB owners, because of Gaylord Broadcasting's ownership of cable TV stations.

[60] Other owners of the period matched Pohlad, Jacobs, Smulyan and Rose in LBO experience: Jerry Reinsdorf and Eddie Einhorn (White Sox, 1981); Ted Turner (Atlanta Braves, 1976); Taft Broadcasting (Philadelphia Phillies, 1981–86); Peter Magowan (San Francisco Giants, 1992); Tom Monaghan (Detroit Tigers, 1983–92); Charles Bronfman (Montreal Expos, 1968–90); Walter Haas (Oakland A's, 1980–95).

[61] The Pirates' selling price of $90 million suggests a strict two to one (67/33) equity/debt requirement in the 1990s.

[62] The group was originally headed by Michael Burke, who had run the team for CBS since 1964. Burke, who had met Steinbrenner only a few months before the deal, was forced out within four months after the purchase. Burke was responsible for the price of $10 million for the club that was $3·2 million less than CBS had paid in 1964. By 1976 Steinbrenner had acquired six of the original 11 partnership shares to become 55% majority owner.

the increasingly public nature of the financial operation of major sports franchises, detailed financial information about club ownership remains cloaked in proprietary secrecy. According to theory, MLB franchises during the free agency period have been expanded beyond their profit maxima and are undervalued in accordance with the sportsman and Steinbrenner effects (*B* and *C* in Figure 2). Theory also suggests that the sportsman will sell his club when the financial return from the performance of the franchise, relative to other investments, no longer compensates the sportsman for the on-field performance of his team. Once the decision is made to sell the club the owner would no longer be willing to sacrifice cash flow for winning, and he would operate the club at the talent level that maximizes the value of the franchise. In order to receive the highest possible price for a franchise at the point of sale, a sportsman owner would downgrade the team's talent to the level that maximizes its book value from *B* to *A* in Figure 2. During the period when a franchise is on the market, player costs will be lower and the value of a franchise should be higher than its previous operating value by the precise amount of the *Sportsman effect*.

During the first two decades of free-agency period over 40 MLB franchises were involved in ownership transfers in one form or another, such as partnership buyouts and expansions, public-private consortia and complete ownership transfers. According to the unified theory, the sportsman effect should be uncloaked in the market period prior to each of these transfers, with the possible exceptions of partnership buyouts and sales involving public-private consortia. Preferably, isolation of the sportsman effect would be accomplished through comparisons of actual exchange prices and values of clubs under sportsman operation. Reported exchange prices are reliable to some extent, but they are problematic in that each transaction involves a different bundle of rights, especially rights regarding stadium ownership and lease arrangements. Unfortunately, the exact operating values of sportsman clubs before and after market periods are, for all practical purposes, unknowable.

Recall that the primary source of the franchise undervaluation is the premium paid for playing talent necessary to accommodate the sportsman's preference for winning. Reliable salary information is available for all clubs during each year of the free-agency period. According to theory, sportsman owners will unilaterally reduce player costs in an attempt to make a franchise more marketable and receive the maximum price. Isolation of the sportsman effect can then be accomplished through the estimation of a team-salary equation over the free-agency period with a binary variable θ for the marketing period preceding each of the ownership transfers:[63]

$$\underset{}{\ln S} = \underset{(62{\cdot}85)}{5{\cdot}24} + \underset{(12{\cdot}67)}{1{\cdot}28 \ln w} + \underset{(6{\cdot}44)}{{\cdot}165 \ln p} + \underset{(62{\cdot}48)}{{\cdot}149\, t} - \underset{(-6{\cdot}20)}{{\cdot}262\, \theta} \qquad \underset{}{R^2 = {\cdot}898} \quad \underset{}{N = 494}$$

[63] Twenty-six MLB ownership transfers, in which the clubs were marketed wholly within the free agency period, are listed in Vrooman (1997a) along with the period the clubs were on the market and the reported exchange prices. Partnership expansions (5), complete ownership transfers (13) and sales following the death (4) or financial distress (4) of the owner are included in the sample, and partnership buyouts (11) and public-private transactions (2) are excluded.

where S is the average salary for team i in period t, w is the average winning percentage for team i in periods t and $t-1$, p is a club's home market population in 1990, and $\theta = 1$ for each of the years that a club was on the market.[64] The existence of the sportsman effect is confirmed by the significance of the θ coefficient at ·01. These estimates, which control for league salary escalation, market size and talent-cycle phase, suggest that player costs were reduced by at least one-fourth during the periods in which these MLB franchises were for sale.

IV Conclusion

Summary. The general theory developed in the first section of this paper leads to three 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, revenue sharing among teams paradoxically does not affect competitive balance among teams, but, it leads to the exploitation of players. The extensive revenue-sharing in the NFL leads to the depression of player salaries. According to the general theory, the NFL is probably the most exploitative 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 in the NBA over the period studied, the NBA is virtually 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 over the period studied. 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.

Labour strife will continue in MLB as long as the baseball players labour market is segmented by lower-tier monopsonistic exploitation and upper-tier monopolistic inefficiency. There is evidence, in this analysis, that free agent talent predictably migrates from quality small market teams to franchises in large markets, but there is also evidence that such movement does not result in large-market dominance. Superior teams are dismantled through the inefficient free agent acquisitions of large-market clubs. This leads to the conclusion that

[64] Student 't' ratios are shown in parentheses The model was estimated for the 26 MLB clubs in existence for the free-agency period 1976–95. The 1993 expansion Florida Marlins and Colorado Rockies were excluded. With the exception of θ, these results for the entire period are consistent with those discussed above (Vrooman, 1995) for the post-collusion period 1990–92: $\ln S = 4{\cdot}39 + 1{\cdot}304 \ln w + {\cdot}223 \ln p + {\cdot}140t$.

the reallocation of free-agent talent under its current constraints is a negative-sum game, and that the inefficiencies originating from arbitrary free-agency criteria have contributed to the inter-seasonal performance discontinuity that reflects an inferior submaximal league product. There is little support in the literature that the uncertainty of a game's or season's outcome, achieved at the expense of total league product, is preferred by fans. Equally bad teams beating one another is an uncertain outcome derived from an inferior league product. As argued in this paper, this is the flawed result of the current limitations placed on free-agent eligibility.

Throughout the era of free agency MLB has experienced a cost squeeze due to the arbitrary segmentation of its labour market. The continuation of salary arbitration and the proposed imposition of cost-sharing salary caps or luxury taxes would continue to deepen player exploitation and serve to reinforce the revenue advantages of large-market clubs in MLB. The monopolistic over-payment of free-agent talent and the exploitation of talent under the reserve clause can be simultaneously reduced if the pool of players eligible for free agency is allowed to increase (and the number of reserved players to correspondingly decrease) to an optimal level. The specific free-agency fix proposed for MLB suggests that salary arbitration for players be eliminated, and that the eligibility requirement for free agency be set at four years, rather than the current six years of MLB service. This would allow MLB clubs to recoup player development expenses, while reducing the number of reserved players by almost one-fourth and increasing competition in the free agency pool by approximately 50%.

The unified theory developed in this paper explores the implications of the general case of the sportsman owner who seeks to jointly optimize profits and winning. When the sportsman exception becomes the general rule, the financial and ownership structures of a sports club necessarily affect the demand for athletic talent. Specifically, the use of financial leverage or ownership syndication creates an incentive for sportsman owners to expand the talent level of their clubs beyond the optimum that would prevail if they were bearing the full cost alone. These *Steinbrenner* agency effects within the environment of restricted free agency are a primary source of the salary escalation in MLB during the free agency period, rather than free agency itself. The theory suggests that, under league competition, the *Steinbrenner effects* from leveraged syndication will increase player costs and reduce franchise values for owners of all clubs within the league. At critical concentrations of syndicated Steinbrenner clones within the league sole unlevered sportsmen will be driven into extinction, and highly leveraged ownerships will be ultimately forced into insolvency. The *Steinbrenner* agency effects are necessarily dependent upon the Sportsman effect, the existence of which is confirmed by evidence that player costs are cut by one-fourth as owners revert from sportsmen to value maximizers immediately prior to selling their clubs.

Directions for further research. Over the course of the twentieth century the major issues in American sports leagues have evolved from a single league's monopsony power over players to include the monopoly power of several

leagues over their fans. Current research on the economics of American professional sports has turned to the question of public provision of sport venues for private gain. Monopoly sports leagues have actively engaged their home markets in political games of stadium extortion through the threat of franchise relocation (Vrooman, 1997b). If a league shares revenue extensively among clubs (NFL) then the league's extortion threats are credible. As the share of revenue retained by the club increases (MLB, NBA and NHL) then extortion games gradually become idle threats made by individual clubs. During the free-agency period, over one-half of NFL franchises have played extortion games with their home markets and one-half of those threats were realized. By comparison, MLB clubs have threatened to abandon their homes to leverage public stadium concessions on several occasions, but not one franchise has actually moved since the Washington Senators became the Texas Rangers in 1972. Empirical evidence stands unanimously negative against the public benefits of sports venues (Noll and Zimbalist, 1997). The community of academic economists universally holds that public provision of stadia is a socially inefficient result of the monopoly power of sports leagues to determine franchise location and league size. Over this century, American professional sports leagues have emerged as double-threats to competition. At the beginning of the twentieth century, a lone monopsony league had sought to 'remedy the evil' of its own labour-market competition by imposing the reserve clause. By the dawn of the twenty-first, monopoly leagues are seeking to further 'remedy' competition by extorting public funds for private gain.

REFERENCES

AKERLOF, G. A. (1970). The market for 'lemons': quality uncertainty and the market mechanism. *Quarterly Journal of Economics*, August, pp. 488–500.

BALFOUR, A. and PORTER, P.K. (1991). The reserve clause in professional sports: legality and effect on competitive balance. *Labor Law Journal*, January, pp. 8–17.

BLASS, A. A. (1992). Does the baseball labor market contradict the human capital model of investment? *The Review of Economics and Statistics*, May, pp. 261–8.

BRUGGINK, T. H. and ROSE, D. R. (1990). Financial restraint in the free agent labor market for major league baseball: players look at strike three. *Southern Economic Journal*, April, pp. 1029–43.

CANES, M. E. (1974). The social benefits of restrictions on team quality. In R. Noll (ed.), *Government and the Sports Business*. Washington: Brookings.

CASSING, J. and DOUGLAS, R. (1980). Implications of the auction mechanism in baseball's free agent draft. *Southern Economic Journal*, July, pp. 110–21.

COASE, R. (1960). The problem of social cost. *Journal of Law and Economics*, October, pp. 1–44.

CYMROT, D. J. (1983). Migration trends and earnings of free agents in major league baseball, 1976–79. *Economic Inquiry*, October, pp. 545–56.

DALY, G. (1992). The baseball player's labor market revisited. In P. Sommers (ed.), *Diamonds Are Forever: The Business of Baseball*. Washington: Brookings, pp. 11–28.

DALY, G. and MOORE, W. J. (1981). Externalities property rights and the allocation of resources in major league baseball. *Economic Inquiry*, January, pp. 77–95.

DEMMERT, H. G. (1973). *The Economics of Professional Team Sports*. Lexington, MA: D.C. Heath.

DEMSETZ, H. (1972). When does the rule of liability matter? *Journal of Legal Studies*, January, pp. 13–28.

EL HODIRI, M. and QUIRK, J. (1971). An economic model of a professional sports league. *Journal of Political Economy*, November/December, pp. 1302–19.
FORT, R. and QUIRK, J. (1995). Cross-subsidization, incentives, and outcomes in professional team sports leagues. *Journal of Economic Literature*, September, pp. 1265–99.
HUNT, J. and KENNETH, L. (1976). Dominance, recontracting and the reserve clause: major league baseball. *American Economic Review*, December, pp. 936–43.
JENSEN, M. and MECKLING, W. (1976). Theory of the firm: managerial behavior, agency costs and ownership structure. *Journal of Financial Economics*, pp. 305–60.
KRAUTMANN, A. C. (1990). Shirking or stochastic productivity in major league baseball? *Southern Economic Journal*, April, pp. 961–68.
LEHN, K. (1990a). Property rights, risk sharing and player disability in major league baseball. In B. Goff and R. Tollison (eds.), *Sportometrics*. College Station: Texas A&M Press.
LEHN, K. (1990b). Information asymmetries in baseball's free agent market. In B. Goff and R. Tollison (eds.), *Sportometrics*. College Station: Texas A&M Press.
MARKHAM, J. and TEPLITZ, P. (1981). *Baseball Economics and Public Policy*. Lexington, MA: D.C. Heath.
MODIGLIANI, F. and MILLER, M. (1958). The cost of capital, corporation finance and the theory of investment. *American Economic Review*, June, pp. 261–97.
NEALE, W. C. (1964). The peculiar economics of professional sports. *Quarterly Journal of Economics*, February, pp. 1–14.
NOLL, R. G. (1974). Attendance and price setting. In R. Noll (ed.), *Government and the Sports Business*. Washington: Brookings.
NOLL, R. G. (1991). Professional basketball: economic and business perspectives. In P. Staudohar and J. Mangan (ed.), *The Business of Professional Sports*. Urbana: University of Illinois Press.
NOLL, R. G. (1995). The economics of sports leagues. In G. Uberstine (ed.), *Law of Professional and Amateur Sports*. New York: Clark, Boardman and Callaghan.
NOLL, R. G. and ZIMBALIST, A. (1997). Build the stadium—create the jobs! In R. Noll and A. Zimbalist (eds.), *Sports Jobs and Taxes: The Economic Impact of Sports Teams and Stadiums*. Washington: Brookings.
PORTER, P. K. (1992). The role of the fan in professional baseball. In P. Sommers (ed.), *Diamonds are Forever: The Business of Baseball*. Washington: Brookings.
PORTER, P. and SCULLY, G. W. (1990). Measuring managerial and allocative efficiency: the case of baseball. In B. Goff and R. Tollison (eds.), *Sportometrics*. College Station: Texas A&M Press.
QUIRK, J. and EL HODIRI, M. (1974). The economic theory of a professional sports leagues. In R. Noll (ed.), *Government and the Sports Business*. Washington: Brookings.
QUIRK, J. and FORT, R. D. (1992). *Pay Dirt: The Business of Professional Team Sports*. Princeton: Princeton University Press.
RAIMONDO, H. (1983). Free agents' impact on the labor market for baseball players. *Journal of Labor Research*, Spring, pp. 183–93.
ROTTENBERG, S. (1956). The baseball players labor market. *Journal of Political Economy*, June, pp. 242–58.
SCITOVSZKY, T. (1943). A note on profit maximisation and its implications. *Review of Economic Studies*, pp. 57–60.
SCULLY, G. W. (1974). Pay and performance in major league baseball. *American Economic Review*, December, pp. 915–30.
SCULLY, G. (1989). *The Business of Major League Baseball*. Chicago: University of Chicago Press.
SCULLY, G. (1995). *The Market Structure of Sports*. Chicago: University of Chicago Press.
SOMMERS, P. M. and QUINTON, N. (1982). Pay and performance in baseball: the case of the first family of free agents. *Journal of Human Resources*, Summer, pp. 426–36.
VROOMAN, J. (1995). A general theory of professional sports leagues. *Southern Economic Journal*, April, pp. 971–90.
VROOMAN, J. (1996). The baseball players labor market reconsidered. *Southern Economic Journal*, October, pp. 339–60.
VROOMAN, J. (1997a). A unified theory of capital and labor markets in Major League Baseball. *Southern Economic Journal*, January, pp. 594–619.

VROOMAN, J. (1997b). Franchise free agency in professional sports leagues. *Southern Economic Journal*, July, pp. 191–219.
ZIMBALIST, A. (1992a). *Baseball and Billions*. New York: Basic Books.
ZIMBALIST, A. (1992b). Salaries and performance: beyond the Scully model. In P. Sommers (ed.), *Diamonds Are Forever: The Business of Baseball*. Washington: Brookings.

Date of receipt of final manuscript: 14 March 2000.