COMPETITION LAW AS A CONSTRAINT ON MONOPOLISTIC EXPLOITATION BY SPORTS LEAGUES AND CLUBS

STEPHEN F. ROSS
University of Illinois

The sports industry is characterized by dominant leagues and clubs exercising economic power unconstrained by rivals or the threat of entry, often featuring market-division schemes. Leagues and clubs can raise price, lower output, and lower quality to fans, create an artificial scarcity of top-tier teams resulting in publicly subsidized stadiums, and impose labour-market restraints that significantly harm consumers by misallocating players, most obviously by inhibiting low-quality teams’ quick improvement. Business decisions made by club-run leagues feature significant transaction costs, resulting in even greater inefficiency than would occur if leagues were controlled by a single entity. Many countries have employed settled principles of competition law, originating in the common law of restraint of trade, as a useful and meaningful constraint on the abuses of economic power in sports. Courts have prohibited agreements between clubs or leagues that distort prices or output, or render output unresponsive to consumer demand, unless the agreement is shown to be demonstrably necessary to achieve a pro-competitive goal. In this paper, I argue that consumers and sports fans will benefit from a more ambitious enforcement of these established principles of competition law.

I. INTRODUCTION

In many economically significant markets within the ‘sports industry’, dominant leagues and clubs exercise economic power unconstrained by rivals or the threat of entry. This exercise of power has predictable results, including an inefficient allocation of resources, often featuring sub-optimal product quality, and significant consumer exploitation. Shielded by entry barriers and horizontal market-division schemes, leagues and clubs can, absent governmental regulation, profitably raise price and lower quality to sports fans. Entry barriers can similarly create an artificial scarcity of top-tier teams, bestowing the power to demand publicly subsidized stadiums and below-market stadium rentals. Collective selling or market division regarding television rights can reduce output and raise prices. Finally,
labour-market controls may unnecessarily restrain wages and freedom of choice for players. These restraints can also significantly harm consumers by misallocating players among teams in a suboptimal way, most obviously by the inhibition of low-quality teams’ ability to quickly improve. Most significantly, because many business decisions are made by club-run leagues, significant transaction costs result in even greater inefficiency than would occur if leagues were controlled by a single entity.

There is a variety of policy alternatives to the distorting and exploitive effect of the exercise of economic power in sports. We could do nothing; government could directly intervene through active regulation; government could alternatively require a restructuring, resulting in at least two competing leagues in each country in each sport. The direction that most countries seem to be moving in, however, is to use settled principles of competition law, which originate in the English common law of restraint of trade, as a useful and meaningful constraint on the abuses of economic power in sports. Whether arising under the common law, the American Sherman Act, or the EU Treaty, courts have prohibited agreements between clubs or leagues that distort prices or output, or render output unresponsive to consumer demand, unless the agreement is shown to be demonstrably necessary to achieve a pro-competitive goal.

In this paper, I argue that consumers and sports fans will benefit from a more ambitious enforcement of these established competition-law principles. Part II identifies several markets where sports products and services are bought and sold and explains why entry or close substitutes rarely constrain incumbents. Absent government intervention, these markets will, therefore, feature the significant exploitation of market power by dominant leagues and clubs. Part III briefly discusses different policy responses to this exploitation, and details how the law’s requirement that competition-lessening agreements be justified provides a workable and meaningful constraint on abuses of economic power in sports markets. Part IV applies these competition-law principles to contemporary agreements posing significant threats of monopolistic or monopsonistic abuse.

II. IN RELEVANT SPORTS MARKETS, INCUMBENTS OFTEN FACE NEITHER CLOSE SUBSTITUTES NOR CREDIBLE THREATS OF ENTRY

There are four principal markets in which sporting products are bought and sold: (i) individual clubs offer admission to a live sporting event; (ii) clubs purchase or rent stadium facilities; (iii) the rights to televise sporting events are sold, individually by clubs or collectively by leagues, to free-to-air broadcasters or pay-television programmers; and (iv) clubs, under regulations determined by the league, acquire the services of skilled players. For many fans, there are no close substitutes to tickets for admission to a game of their favourite team. For most fans, matches from other sports are not close substitutes. Thus, individual teams have successfully developed sufficient product differentiation and brand loyalty to set price and quality at a level unconstrained by rivals or entry.

On the margin, there are current or potential consumers who do not have loyalty. Because it is difficult to price-discriminate between elastic and inelastic demanders, many teams price tickets without regard to those marginal consumers. However, entry might promote consumer choice and provide an alternative for these marginal consumers, at least in geographic markets large enough to permit such entry at minimum viable scale. North American clubs have agreed to exclusive territories, precluding such entry. Thus, while some London football fans might switch allegiance, or new fans might have been drawn in when Fulham was promoted several years ago, while West Ham United will significantly suffer from its relegation next season, football and hockey fans in Chicago have no reasonable substitutes to the mediocre Bears and Blackhawks clubs.

In a competitive market, consumers will find sellers willing to provide the demanded output at a price covering costs, including the normal cost of essential facilities such as a stadium. Sports leagues with entry barriers, however, are not constrained by the competitive process in this regard. By limiting the number of teams within their league, member clubs
can create a scarcity that results in consumers, through their local governmental authorities, bidding for scarce franchises by constructing stadiums at taxpayers’ expense and through generous rental agreements. This phenomenon is a frequent occurrence in the United States and rare in Europe. The threat of entry through the promotion and relegation process featured in European sports primarily accounts for this significant comparative difference.

Brand loyalty and differentiated products present related issues with regard to the demand for watching matches on television. Clubs and leagues seem to follow different strategies, with some seeking fully to exploit inelastic demanders by selling rights to pay cable, while others appeal to broader audiences on free-to-air telecasts. Shifts from free-to-air to pay not only result in substantial output reduction (estimates are that English football viewership declined from 7m to 1m when games shifted from free-to-air to BSkyB), but also a significant wealth transfer from fans who previously enjoyed a huge consumer surplus from free-to-air. In light of the fierce loyalty many football fans show to their favourite club, agreements that leave substantial numbers of games untelevised create substantial welfare losses.

Although fans may have strong loyalty to their favourite club, they may see major games among other teams in the league as close substitutes for each other. Similarly, in tiered structures prevalent in Europe, some fans may maintain their prime loyalty for lower-tier clubs and view various telecasts among top-tier clubs as equivalent. However, competition in this market is often limited by agreement among the clubs. Most prevalent is the practice of collective selling, so that clubs do not compete with each other at all in the sale of television rights. Even where, as is prevalent in North American sports other than football, clubs retain rights to sell games individually, competition is significantly constrained by exclusive territorial agreements that bar teams from selling their games outside a defined local market.

The lack of substitutes to constrain the exercise of power by sports clubs is particularly problematic because rights sales are made by leagues governed by the member clubs whose primary interest is their own individual welfare, not the league’s overall welfare. The experience of the Premier League demonstrates that, as technology expands opportunities for rights sales (through broadband or increased number of channels available on satellite and cable television), solutions that are output- and profit-enhancing may not be selected by clubs because transactions costs prevent the parties from agreeing on how to share the increased revenue (Szymanski and Ross, 2000).

Clubs traditionally faced little constraint on the exercise of monopsony power in labour markets, because of pervasive agreements among all members of the league (or, in the case of soccer, international agreements organized by FIFA). These restraints have little effect on supply because, with the exception of the two codes of rugby, players at the highest level of sport rarely possess the ability to switch sports, and the next-best occupation for almost all players usually pays far less than even the monopsony price for their labour. Restraints may affect overall quality, in terms of training and fitness of athletes, however. Regardless of the effect on supply, these restraints result in huge wealth transfers from workers to clubs.

No serious allocation problems exist where, as in European football, player contracts are freely transferred for substantial sums reflecting market value. However, because of a desire to avoid the public-relations and collective-bargaining problems that would be created by such a transparent gap between players’ salaries and their true value (as reflected in a transfer fee), and as one means to promote competitive balance, North American and Australian leagues have traditionally inhibited player sales. This results in an artificial rigidity in the labour market that leads to inefficient allocation of players, most particularly in the inability of inferior teams to improve quickly.

In sum, leagues and clubs operate in a number of relevant economic markets. In many cases, entry barriers and product differentiation preclude a market constraint on the ability of leagues and clubs to exercise economic power.
III. COMPETITION REGULATION PROVIDES A FAIRLY UNIFORM AND MEANINGFUL CONSTRAINT ON INEFFICIENT OR EXPLOITIVE ABUSES OF POWER BY SPORTS LEAGUES AND CLUBS

(i) Policy Alternatives

If the normally self-correcting features of economic markets are unlikely to constrain the exercise of economic power by sports leagues and clubs, what policy alternatives are there? A variety of approaches are possible, including *laissez faire*, direct government regulation, or indirect government regulation by assuring the existence of multiple competing leagues in each sport. As detailed below, each of these raises significant economic and political problems. Operating through private and governmental enforcement of competition statutes and the common law on restraint of trade, however, courts have had modest success in ameliorating inefficient or exploitive abuses of power. More vigorous enforcement through the courts offers an excellent opportunity for greater protection of the public interest.

One could adopt the view that any governmental ‘cure’ through intervention in the free market is likely to be worse than the disease, and opt for a complete *laissez-faire* approach to sports. Although monopolistic abuses in sport, to be sure, raise less of a societal concern than when products essential to the economy are subject to output restraints or inefficiencies, huge social-welfare costs still exist. Without threat of entry, North American clubs receive billions of dollars in tax subsidies for stadiums, while many are deprived of high-level competition by the closed league structure there; freed from any governmental restraint, European football clubs may well follow. Unrestrained ability collectively to sell television rights results in significant reductions of output, wealth transfers from consumers, and distorting effects in the more general market for television. Absent legal restrictions on employer conspiracies in labour markets, clubs will distort the efficient allocation of players. Moreover, as evidenced in the United States by the record of Major League Baseball (which historically enjoyed an exemption from competition-law regulation), employer monopsonies lead to significant industrial unrest in the form of player strikes and lock-outs as employers seek to maintain their power.

Direct governmental regulation of potentially abusive practices is also undesirable. As has been previously detailed (Ross, 1989, pp. 702–14), the tasks involved do not lend themselves well to governmental administration. Standards for determining the precise number of teams in a closed-structure league, or of the proper allocation of franchises in such a league, are extremely difficult to formulate. Determining proper labour-market rules that combine the public interest in efficient allocation of players with the stronger public interest in industrial peace, in a format outside of collective bargaining, is also extremely problematic. As evidenced by the more modest controversies over listed events legislation in Europe and Australia, complete regulation of the terms of sale of all television rights to sporting events seems hopelessly complex. Moreover, even if brilliant bureaucrats could succeed at command-and-control regulation, the sports industry is ripe for regulatory capture. Regulation principally constrains a few major sports businesses with huge incentives for governmental lobbying, and generally benefits the widely dispersed interests of sports fans and taxpayers. This is not a good recipe for regulatory success (Olson, 1971).

Another form of legal intervention would allow firms mostly to do as they pleased, provided that each sport had at least two vigorous rival competitive leagues (inter-league competition for a championship would be permitted but closely scrutinized). The advantage of such a system is that it requires minimal governmental intervention: the supervision of a divestiture, and the policing of conduct solely to prevent predatory conduct and all but a small number of inter-league agreements deemed essential for championships. Output-reducing limits on the number and location of franchises would be quickly checked by the rival league. Output-reducing television agreements would be severely constrained by the ability of the other league to increase telecasts. Monopsony power in labour markets would not be possible, so restraints would be designed for efficient player allocation.

There are, however, significant legal and economic obstacles to such an approach. Current conduct by
dominant sports leagues, while clearly reflecting the exercise of monopoly power, may not meet the precise standards set forth in statutory or case law to justify divestiture. Divestitures may also be inconsistent with a ‘European model of sport’ that features national federations as the dominant regulator of competition within a sport (the most plausible result of a divestiture would be two or more pan-European football leagues, for example, replacing the dominant domestic leagues). Efforts to obtain a divestiture in the United States would be further hampered by a judicially created antitrust exemption for baseball, and a specific statute that allows football leagues to merge. Others question whether long-term competition between rival leagues is viable. Although history reveals that competition between rival North American leagues have inevitably ended owing to merger, predatory conduct, or imprudent business decisions (Ross, 1989), some commentators have suggested that a long-term competitive equilibrium is not realistic and that the desire for a superior league will inexorably result in a single league’s dominance once it is perceived as superior, thus allowing the leader to regain monopoly power (Roberts, 2003).

Thus, it would appear that the most practical option for constraining monopoly power would be to subject agreements among the dominant league’s clubs to a standard of reasonableness. This standard requires that agreements which appear to have anticompetitive effects must be justified as reasonably necessary to achieve some legitimate welfare- or consumer-enhancing goal. This standard evolved from the English common law of restraint of trade, was further developed by judicial precedents interpreting the American Sherman Act to incorporate common-law concepts, and has now been effectively adopted as the governing standard for applying, in the sports context, Section 1 of the American Sherman Act and Article 82 of the EC Treaty.

(ii) The Common Law of Restraint of Trade

(The following discussion summarizes a forthcoming law-review article detailing the application of the common law of restraint of trade to sports restraints.)

Although early common-law decisions, usually condemning restraints applied involuntarily by force of law, royal prerogative, or custom, seemed to proscribe any effort to ‘restrain any to use a lawful trade at any time or at any place’ (Colgate v. Bachelet (1601)), the landmark decision in Mitchel v. Reynolds (1711) established that some voluntary restraints were lawful. That decision also established a bedrock principle of particular relevance to professional sports-league trade restraints: restraints broader than necessary to protect legitimate interests are invalid. This principle was restated in what remains the leading restraint-of-trade case in the Commonwealth, Nordenfelt v. Maxim Nordenfelt Gun & Ammunition Co. (1894), where (pp. 554–5) Lord Macnaghten defined the reasonableness standard thus:

> reasonable, that is, in reference to the interests of the parties concerned and reasonable in reference to the interests of the public, so framed and so guarded as to afford adequate protection to the party in whose favour it is imposed, while at the same time it is in no way injurious to the public.

This test was further refined in Kores Mfg Co. v. Kolok Mfg Co. (1958), where the English Court of Appeal explicitly required that labour restraints be justified by legitimate employer interests, and specifically held (p. 125) that employers have no such interest ‘in preventing an employee, after leaving his service, from entering the service of another competitor merely on the ground that the new employer is a competitor’. As Lord Diplock explained in Petrofina (Great Britain) Ltd v. Martin (1966, p. 182), the public interest determines what rights are legitimate. Years earlier, the House of Lords announced that ‘freedom from competition per se apart from [legitimate interests], however lucrative it might be’, was not entitled to protection (Herbert Morris Ltd v. Saxelby (1916, p. 702)).

The second important tool that common-law judges use to protect the public interest in competition is to require that restraints be narrowly tailored to achieve legitimate interests. For example, in Mineral Water Bottle Exchange and Trade Protection Soc’y v. Booth (1887), the English Court of Appeal denied an injunction to enforce a trade association’s rule preventing the hiring of employees of another member for 2 years. Acknowledging that employers had a legitimate interest in preventing the misuse of confidential information, the court reasoned that the restraint applied indiscriminately to those who had never been in a confidential position nor posed any...
threat to misuse information acquired during their previous employment.

Thus, when this doctrine is applied in the context of professional sports leagues, courts have held that agreements that strictly limit the ability of players to obtain competitive bids for their services are illegal under the common law. As made clear by the New Zealand Court of Appeal in Blackler v. New Zealand Rugby Football League (1968, p. 571), restraints must be justified as legitimate in light of public-interest considerations that include consumer interests and individual freedom for players. The pathmarking case is Eastham v. Newcastle United Football Club (1963), a challenge to the English ‘retain and transfer’ rule, similar to baseball’s ‘reserve clause’, which bound a player to his former club, even if the contract had expired and the club was not paying him! Finding a real inequality in bargaining power, Lord Wilberforce stated (p. 438) that the pervasive use of this system by sports-league employers justified careful judicial consideration of whether the rules went ‘further than is reasonably necessary to protect their legitimate interests’. He recognized that the Football League had a special and legitimate interest in maintaining the overall quality of the sport through competitive balance. He acknowledged (pp. 431–2) that, if richer teams could acquire most of the better players, this would be ‘to the detriment of the whole’. The burden that Eastham places on owners and leagues cannot be overemphasized. Lord Wilberforce’s opinion demands proof not only that richer teams would, indeed, be more active in signing free agents than financially poorer clubs, but that the free agents signed would be better, thus directly harming competitive balance. His Lordship correctly identified, however, several reasons why the league’s case failed to demonstrate the need for these restrictions. First, he noted (p. 433) that the clubs commonly transferred rights to players to other clubs for a fee, and that richer clubs already used their superior resources to buy players. Second, he noted that clubs with good employee relations could draw upon the natural reluctance of a player to betray feelings of loyalty as well as to disrupt his family. Third, he also observed (p. 434) that ‘[n]o club desires more than so many centre forwards’, so that players are unlikely to switch to talent-laden teams where they might not make the first eleven.

Thus, he concluded (p. 436) that the evidence did not support the ‘rather far-fetched argument’ that the rule tended to spread good players evenly over the various clubs in each division.

Lord Wilberforce’s third point deserves additional note. All things being equal, a player is, indeed, more likely to be valued more highly by an inferior team, and so the free market naturally tends to improve competitive balance. A championship team will find it very difficult to sustain the salary demands of all its players; a team with a disappointing record will have a greater incentive to demonstrate improvement next year, and is more likely to find players with subordinate roles on top teams who can become stars on their own rosters. Both American and Australian commentators have observed that unrestrained labour markets permit players to move from superior to inferior teams and restrained markets often rigidify the dominance of a few teams (Ward, 1985, p. 546; Ross and Lucke, 1997).

Two other common Justifications for sports league restraints were also rejected in Eastham. Lord Wilberforce found (p. 435) that long-term stability was not a legitimate basis for restraints, because clubs could reduce turnover by signing players to staggered long-term contracts. He also rejected (pp. 435–6) the claim that the restraints were necessary to secure clubs’ investment in training and development of young players. The answer, again, lay in individually negotiated terms in contracts with younger players to ensure that the club had a reasonable opportunity to recoup its investment. For example, the court in Toronto Marlboro Major Junior ‘A’ Hockey Club v. Tonelli (1975), described contract language signed by teenage hockey players requiring payment of $25,000 to the league upon signing a major league contract. Likewise, although Greig v. Insole (1978, p. 499) suggested that a dominant league has a legitimate interest in preventing free riding from a new entrant who ‘is creaming off . . . the star players who it has itself incurred no expense in training and preparing for stardom’, this legitimate concern can be adequately addressed by some individually negotiated contract between younger players and national organizations (Trebilcock, 1986, p. 229).

Eastham was followed by the Australian High Court in Buckley v. Tutty (1971). Buckley agreed
(p. 377) that the leagues had a legitimate interest in ensuring that ‘teams fielded in the competitions are as strong and well matched as possible, for in that way the support of the public will be attracted and maintained, and players will be afforded the best opportunity of developing and displaying their skill’. Note the important gloss that the Court placed on Eastham’s validation of competitive balance — it is not an end in itself but only a means to the ends of maximizing fans’ interest and the most efficient allocation of player resources.

It is no doubt true, as Lord Wilberforce expressly held in Eastham (p. 148), that ‘[r]egard must be had to the special character in which the restraints operate’ in the sporting context. The judge in Greig v. Insole (p. 497) held that national sporting organizations may have a bit more leeway to restrain trade if ‘reasonably necessary for the protection of the organisation and administration of the game’. Such an open-ended standard meant that the floodgates have been opened to an enormous volume of evidence in regard to the organisation of first-class cricket, the structure and finances of its organising bodies and the lives of its players in this country and many other countries. (ibid.)

However, ‘it is difficult to see how some sort of factual inquiry about actual or likely impacts of restrictive policies can be avoided if the public interest test is to be given a coherent and meaningful context’ (Trebilcock, 1996, p. 220). Significantly, Greig (pp. 502–3) found that even significant financial losses by a national federation which used the revenues to promote the sport on a grassroots level did not outweigh the impact on players and the deprivation to the public of watching the highest quality of the sport possible. The judge (p. 500) carefully parsed the specifics of the case to determine the precise extent that a restraint was necessary to preserve the game.

Building on restraint-of-trade precedents protecting an individual’s right to work and applying the general test of Nordenfelt, an unreported Chancery decision also suggests the applicability of this doctrine to judicial oversight of potential abuses in the ‘pyramid’ structure of football and other European sports, where clubs’ ability to participate in competitions is subject to their ability to be promoted to higher levels of competition or relegated to lower levels. In Stevenage Borough Football Club Ltd v The Football League Ltd (1996), the court found that rules that unreasonably denied clubs the opportunity for promotion were invalid as a restraint of trade. In that case, the court took issue with a Football League rule that required a club seeking promotion from the lower-tier Football Conference to bring its stadium up to League standards by a date prior to the end of the season, as unfairly requiring a club to make expensive improvements that would only be required if the club actually triumphed during the Conference season. However, because of the strong public interest in the orderly implementation of League rules, the Court held that the onus was on those challenging the rules to show them unreasonable, and further denied relief in this case because the club had waited until the last minute to file a legal challenge to the rules.

In sum, the common law provides a meaningful restraint on welfare-reducing or exploitive conduct. Where a restraint has a significant anticompetitive effect, it must be justified as reasonably necessary to achieve some socially legitimate interest, and the court must balance these interests against those of the parties restrained and the general public.

(iii) The Sherman Act

‘Every contract, combination in the form of trust or otherwise, or conspiracy, in restraint of trade or commerce among the several States, or with foreign nations, is declared to be illegal’ by Section 1 of the Sherman Act. In the landmark United States v. Addyston Pipe & Steel Co. (1898) decision, the court declared (p. 282) that this broad language was intended to make criminal and tortious those contracts that were void as unreasonable restraints of trade under the common law and concluded that the common law tolerated restraints among competitors only when ‘ancillary to the main purpose of a lawful contract, and necessary to protect the covenantee in the enjoyment of the legitimate fruits of the contract’.

As applied to sports, the Supreme Court held in NCAA v. Board of Regents (1984) that the Sherman Act permitted clubs that were otherwise in competition to agree to restrain trade, because some
restraints were necessary in order for the product to exist at all. However, clubs or leagues bore a heavy burden to justify restraints where it appeared that they had the effect of raising price, reducing output, or rendering output unresponsive to consumer demand. Thus, the Court held (pp. 101–2) that colleges could agree to a host of restraints designed to differentiate ‘amateur’ college sports from minor-league professional sports, and a district court in Molinas v. National Basketball Association (1961) held that clubs could agree to boycott players found to be engaged in match-fixing or other corrupt activities. However, NCAA’s reasoning is such that any justification must demonstrate that, at least in the long run, the restraint will result in lower costs or enhanced consumer appeal.

Both the Supreme Court in NCAA and the court of appeals in Mackey v. National Football League (1976) held that competitive balance is a legitimate justification for sports league restraints. However, in both cases the restraints were found overbroad. In NCAA, an output restraint on televised college football games (college football is a major revenue and entertainment source in the United States, with top universities attracting over 100,000 fans to games and programmes in top conferences averaging over $6m in annual profit) was found (p. 119) to be ‘not even arguably tailored’ to any goal of promoting competitive balance among college football programmes. The court also noted that the NCAA (National Collegiate Athletic Association) had failed to pursue other alternatives—such as revenue sharing or limits on expenditures—that might achieve the goal. In Mackey, the court found a rule authorizing the league commissioner to require significant ‘compensation’ if a player switched employers, to be similarly overbroad in restricting stars and journeymen, regardless of whether the transfer was to an already dominant team or the reverse.

A similar approach, pre-dating Supreme Court guidance, was used by the district court in United States v. National Football League (1953). The court held that the league could reasonably agree on a rule that clubs would not televise games into the home territory of other clubs at the same time the home team was playing. Such a rule was found to be reasonably necessary to protect the live gate of home clubs, which was essential to the home team’s viability, and thus the league’s success. However, agreements that clubs would not broadcast, via television or radio, games into another’s home territory at the same time the home team was broadcasting its own games, was held to be unreasonable. The effect of such a rule was simply to reduce output and to limit competition to permit price increases.

Like common-law courts, American antitrust decisions do recognize that leagues and clubs have a legitimate interest in facilitating desirable investment by preventing free riding on others’ efforts. However, when the ‘reasonably necessary’ principle of the common law and Addyston Pipe is applied, it is difficult for an intra-league sports restraint to be justified on free-rider grounds, for the obvious alternative of sharing revenue is almost always available. As the court noted in Chicago Professional Sports Ltd Partnership (1992), in rejecting an effort by the National Basketball Association (NBA) to limit the number of Chicago Bulls games featuring Michael Jordan from being shown on national cable television:

What gives this the name free-riding is the lack of charge. . . When payment is possible, free-riding is not a problem because the ‘ride’ is not free. Here lies the flaw in the NBA’s story. It may (and does) charge members for value delivered. As the NBA itself emphasizes, there are substantial revenue transfers, propping up the weaker clubs in order to promote vigorous competition on the court. Without skipping a beat the NBA may change these payments to charge for the Bulls’ ride. (p. 674)

American antitrust law may depart from common-law doctrine in one respect, although the actual implications for sport are unclear. Addyston Pipe held that restraints must be justified as necessary to permit economically integrated parties (partners, joint ventures, sellers, and buyers) to promote competition in some manner. It acknowledged that this view—now widely accepted in the United States—was inconsistent with several late-nineteenth-century English and Canadian common-law opinions that appeared to permit restraints justified by the need to serve some undefined public interest or prevent ‘ruinous competition’. The court criticized these opinions, declaring (p. 284) that they ‘set sail on a sea of doubt’, permitting judges to determine ‘how much restraint of competition is in the public interest, and how much is not’. It concluded that the
‘manifest danger in the administration of justice according to so shifting, vague, and indeterminate a standard would seem to be a strong reason against adopting it’.

The potential difference between the Sherman Act and the common law is illustrated by Greig v. Insole’s consideration of the International Cricket Council (ICC)’s justification for blocking a rival cricket league by preventing the league’s players from also participating in ICC test matches. Although, as noted above, the English court found the restriction overbroad, the court declared that the ICC could adopt restraints ‘reasonably necessary for the protection of the organisation and administration of the game’. Under a Sherman Act analysis, a restraint could also be justified if necessary for the viability of the league. In terms of ‘grassroots’ participation, however, the ICC would have to show that the presence of a rival league would cause overall output of cricket to decline or to be rendered unresponsive to consumer demand. The Supreme Court held that the NCAA could not justifiably restrain the output of top college games because of a concern that so many fans would prefer to watch these games that small colleges would lose patronage. Similarly, the risk that a new, consumer-friendly cricket competition draining money from the ICC would result in the threatened viability (through loss of subsidy) of county teams from Northamptonshire and Durban would not justify the restraint under the Sherman Act. It is not clear, however, whether such a justification would be found to be in the public interest under the common law either, especially in light of the weight given by the court in Greig to the public’s interest in viewing quality matches.

There are three recognized and one potential exception to the general policy of the Sherman Act. The Supreme Court has held that baseball is exempt from antitrust analysis. The Court has also held that any agreement that primarily affects a labour market is exempt as long as the players are represented by a union engaged in collective bargaining under the aegis of federal labour law. Congress established a limited immunity for collective selling of television rights. Finally, some commentators and one prominent appellate judge believe that Section 1 should not apply at all to leagues, because they are ‘single entities’ whose inter-club agreements are more akin to group decisions by partners than restraints by independent entities. None of these exceptions deserve emulation elsewhere.

The Supreme Court held in Federal Baseball Club v. National League (1922) that the sport did not constitute ‘trade or commerce among the several States’ that Congress could regulate through the Sherman Act. As a court of appeals recently observed in Major League Baseball v. Crist (2003), the exemption’s initial grounding on a cramped and now-discarded definition of interstate commerce, the Court’s unusual reliance on congressional inaction (in contrast to parliamentary systems with party discipline, the ability of individual legislators or committees to block efforts to overturn unpopular court decisions has usually resulted in a weaker regard for precedent in the United States), the well-documented welfare losses stemming from unreviewable anticompetitive agreements, and the fact that modern antitrust analysis would uphold conduct that is necessary for maintaining a successful league-based enterprise, all meant that the ‘death of the business-of-baseball exemption would likely be met with considerable fanfare’ if the Supreme Court or Congress saw fit to do so.

The National Labor Relations Act provides that, when chosen by a majority of workers in a bargaining unit (a term of art defined by an administrative agency), a union becomes the exclusive representative of all current and potential workers in that unit. The Clayton Act, another antitrust statute, provides that labour is not a ‘commodity or article of commerce’ and exempts trade-union activity from the antitrust laws. Although typical industrial unions seek to reduce competition among workers (by negotiating fixed pay scales, for example), a major priority for players’ unions in the 1970s was to negotiate rules permitting increased competition for player services. The public interest in permitting unions and management to reach compromise agreements that insure industrial stability led courts in McCourt v. California Sports, Inc. (1979) and Wood v. National Basketball Association (1987) to hold that labour-market restrictions agreed to through arm’s-length bargaining, such as limits on competition for veteran players or a cap on total payroll, would be immune to antitrust scrutiny. The Supreme Court unfortunately extended this exemption in Brown v. Pro-Football Inc. (1996) to restrictions not agreed to by the union, thus putting
players to the choice of collectively bargaining or asserting antitrust rights. As detailed elsewhere (Ross and Lucke, 1997), because labour-market restraints can also significantly affect product quality—through the inefficient allocation of players among teams, particularly via restraints that harm competitive balance by making it more difficult for inferior teams to improve—the public interest in competition ought only to be sacrificed when the benefits of labour peace are likely to be achieved.

In 1961, in response to a similar innovation by the rival American Football League, the National Football League entered into a collective selling agreement with a major television network that ensured that each team’s away games would be shown to its local fans, and guaranteed a national market for selected games. Whether or not the overall effect of this arrangement was to reduce viewership (by eliminating the opportunity for individual clubs to sell anywhere) or enhance viewership (by guaranteeing all teams a market for their games plus a nationwide market for some games) was unclear (Ross, 1990, pp. 469–70, n. 28). However, in United States v. National Football League (1961), the court held that the plan violated the literal terms of its order in the 1953 case, and invalidated the deal. The NFL turned to Congress, which passed a specific exemption for any ‘joint agreement . . . by which any league of clubs . . . sells or otherwise transfers all or any part of the rights of such league’s member clubs in the sponsored telecasting of the games’, provided that the purchaser of rights is free to televise games anywhere except within a team’s home territory while that team is playing at home. It was clear that Congress intended this exemption to facilitate what it perceived as a welfare-enhancing deal by the NFL. As such, it may well be unnecessary today. Moreover, the term ‘sponsored telecasting’ has been understood to be limited to free-to-air telecasts, so that the exemption does not apply to sales for pay television.

For years, commentators debated whether Section 1 was appropriately applied to agreements among a league’s clubs, in light of the significant economic integration necessary to promote a league. (References can be found in Ross, 1997, p. 549, n. 136.) Although Courts have frequently rejected this argument, based on the formal organization of leagues as unincorporated associations of separately owned clubs that do not share profits or losses, the independent management of each club, and active competition among clubs that would occur but for challenged restraints (Ross, 1997, p. 550), new life has been breathed into this claim by an opinion of an influential appellate judge and former University of Chicago professor, Frank Easterbrook, in Chicago Professional Sports Ltd v. National Basketball Association (1996).

As detailed elsewhere (Ross, 1997), Judge Easterbrook’s view that leagues are analogous to corporations is severely flawed. As his mentor, Judge Richard Posner, wrote in General Leaseways, Inc. v. National Truck Leasing Association (1984, p. 594), it ‘does not follow that because two firms sometimes have a cooperative relationship there are no competitive gains from forbidding them to cooperate in ways that yield no economies but simply limit competition’. The decision’s analogies reveal the flaws in the reasoning. Judge Easterbrook noted (p. 598) that divisions within a single firm might have competing interests, with separately compensated executives seeking to maximize their own division’s profits, but that these conflicts do not ‘imply that these large firms must justify all of their acts under the Rule of Reason’. The critical difference between General Motors and a sports league, however, is that ultimately the decision whether Saturn should refurbish a run-down Oldsmobile plant (good for Oldsmobile) or build a more expensive plant far away from the traditional culture of Detroit (good for Saturn) will be made either by senior executives or a board of directors concerned only about maximizing revenues for the entire corporation, whereas league decisions are made by the selfinterested owners themselves. Indeed, the inefficiency in allowing clubs to determine league policy has been demonstrating in the context of television sales for English football (Szymanski and Ross, 2000), and recognized by the Australian Football League in creating an independent six-member commission to run league affairs.

(iv) European Community Law

Article 48 of the EC Treaty prohibits restrictions on the free movement of workers through the European Community. Article 81 of the Treaty prohibits
concerted practices which may affect trade between Member States and which have as their object or effect the prevention, restriction or distortion of competition within the common market’. In Union Royale Belge des Sociétés de Football Association v. Bosman (1996), the European Court of Justice found that Article 48 prohibited rules, promulgated by the Union of European Football Associations (UEFA) and adhered to by all European professional football clubs, that required compensation for players signing with new clubs after the expiration of their contract. It did not reach, but by implication its reasoning is consistent with, the prior conclusion of the Advocat General that the rules violated both sections 48 and 81.

In recognizing that the EU Treaty limits the ability of dominant leagues and clubs to restrain trade, Bosman acknowledged (¶104) that restrictions would be compatible with the Treaty if they pursued a ‘legitimate aim’ and were reasonably necessary to achieve the public interest. However, as the Court emphasized in Lehtonen v. Fédération Royale Belge des Sociétés de Basket-ball (2001, ¶56), ‘measures taken by sports federations with a view to ensuring the proper functioning of competitions may not go beyond what is necessary for achieving the aim pursued’. In his preliminary decision in Lehtonen, the Advocat General was more explicit (¶107):

Competition-restricting rules such as in the present case which have the effect of promoting the establishment of competition on the market in question may therefore be compatible with [Article 81] if they are necessary and reasonable for achieving that objective.

The Bosman Advocat General suggested that maintaining competitive balance in leagues and assuring that lower-tier teams were able to recover their investment in developing and training young players were both legitimate aims. The requirement of an agreed-upon or arbitrated transfer fee for all players was, however, an overly broad restriction. Similarly, in Lehtonen the Advocat General recommended invalidating rules limiting late-season roster additions on overbreadth grounds. He conceded that the legitimacy of the competitive season could justify rules that prevented roster additions, but found no basis for requiring that European transfers be consummated by 28 February while players from the United States or Brazil might be added until 31 March. (The Court did not accept the recommendation that the rule be voided, choosing instead to remand the case to the Belgian national courts to determine whether there was an objective justification for the rule.)

Although the basic legal standard applied under European law is, therefore, similar to that of the common law and American antitrust law, its application to Europe’s dominant professional sport—soccer—raises distinct issues. The principal form of North American professional sports competition is club competition within a dominant league. This vibrant form of competition co-exists in Europe with national team competition and the increasingly popular competition among the best clubs within each domestic league for a European club championship. Because top clubs—the ones most likely to be affected by restraints for which competitive balance would be offered as a justification—often play simultaneously in domestic leagues and in European contests, whether the competitive balance actually enhances consumer appeal is unclear. Competitive balance justifies restraints where overall attendance and television viewership increase if each season features a close contest for the league championship, and when each team has a reasonable chance to contend for the championship every few years. Thus, although Ajax, Feyenoord Rotterdam, and PSV Eindhoven consistently dominate the Dutch league, it is not clear if overall output (measured by live attendance plus television viewership) would be enhanced in the Netherlands by greater balance within the Eredivisie. Due to a greater fan base—as reflected in attendance and viewership of games on television—it might be preferable for these clubs to contend a disproportionate amount of time; fan interest may well be less in a competition dominated by Utrecht, William II Tilburg, and Groningen. Fan interest might also be optimized by a system that allowed Ajax consistently to contend for the top position in the Champions League, and that might allow two other Dutch teams, for example, to compete to the final rounds of the UEFA Cup, even though the domestic season is less interesting because of this imbalance.

American and common-law courts have traditionally rejected efforts to justify labour restraints as necessary to promote player development. Mackey v. National Football League (1976) noted that, in
contrast to the uniquely interdependent interest of each club in having other strong, competitive clubs, ensuring that an employer can recoup investment in training young employees is not unique to sports. Legitimate employment interests can be protected by individually contracting with young players. Moreover, leagues keen to promote player development by lower-tier clubs are free to raise and redistribute revenue, for example by a 5 per cent tax on all clubs’ payrolls. Such a undifferentiated tax would not significantly impede bidding for players’ services.

Under Anglo-American law, private restraints to prevent de-stabilizing mid-season contract breaches by players are also unnecessary, because another club seeking to induce such a breach would be liable in tort; often the player can be enjoined from playing for another team; and buy-out clauses reasonably calculated to protect the club’s investment will be enforced. Civil codes generally provide for the enforcement of penalty clauses to assure the performance of freely negotiated long-term contracts (Ross, 1999, p. 106). However, some continental jurisdictions have superseding statutory provisions specifying that workers may breach employment contracts with minimal penalty. This would allow Manchester United to lure players from Ajax in mid-contract, while allowing Liverpool to prevent one of its players jumping to Eindhoven. Focusing on the Treaty’s commitment to free worker movement, European competition authorities take the view that Bosman prohibits restraints on within-contract bidding for services as well as bidding for those, like Bosman himself, whose contractual relationship with his team had expired. Recognizing the legitimate interests involved, a settlement with European football authorities was recently reached, permitting restraints that serve to enforce multi-year contracts, but only for 3 years time.

Although professional sporting institutions in Europe and North America require that core competition principles be applied in different ways, the basic approach is the same: legitimate interests that serve the sport and sporting consumers may justify restraints that are reasonably necessary to achieve their goal. Efforts to achieve goals that are not unique to sport, or efforts to achieve legitimate goals in ways that are overbroad or pretextual, are not permitted.

IV. CURRENT OPPORTUNITIES FOR COMPETITION-LAW INTERVENTION

A variety of current or proposed agreements among leading clubs and leagues may have the effect of raising price, reducing output, or rendering output unresponsive to consumer demand. These anticompetitive agreements are subject to private or governmental challenge based on the Sherman Act, EU Treaty, and the common law.

(i) The Structure of Promotion and Relegation

As demonstrated elsewhere (Ross and Szymanski, 2002), the ability of teams to enter the top tier of domestic competition by proven success at a lower tier has a variety of salutary effects in constraining a dominant league’s opportunities to exploit consumers. Promotion and relegation makes output more responsive to consumer by providing a race-away-from-the-bottom as well as a race-for-the-top that attracts widespread fan interest, provides a strong incentive for teams to invest in player quality, and significantly constrains the ability of incumbent clubs to demand exploitive terms for stadium construction or rental because of the credible ability of a locality interested in a top-tier team to support a lower-tier team’s efforts to improve.

Obviously, the application of these principles to the major North American sports would have the most significant effect. The requirement of promotion and relegation also is an important constraint on reported efforts by leading European football clubs to form their own ‘Super League’ apart from domestic competitions. Although the issues are more complex, the principle also relates to proposed agreements among clubs in the lower-tiered English Football League that would limit the payrolls of these teams. Although motivated by concerns about financial stability, restraints should not effectively limit the ability of a team, with an actual or potential fan base sufficient to compete in the Premier League, to employ players of sufficient talent to obtain promotion.

(ii) Television Contracts

Competition issues relative to the assignment of television rights arise in at least two distinct
contexts. Because sports programming is so desirable, the acquisition of rights to a large number of sporting events may have foreclosing or other effects that harm competition in various television markets. These concerns are beyond the scope of this article. Rather, the focus here is on restraints that raise price, reduce viewership, or render output unresponsive to consumer demand. Rights sales can have this effect principally when rights are sold collectively or when clubs divide markets—usually through agreeing not to sell in each other’s ‘home’ territory.

Collective sales can be output-enhancing in a number of circumstances. Granting a programmer exclusive rights to telecast prime matches deliberately scheduled at atypical times (Monday Night Football) is one example. A sale that ensures a nationwide outlet for a league is another example. Many thought that the exclusive sale of Premier League games in 1992 to BSkyB, resulting in increased output over the status quo, new technology, and the facilitation of a new entrant into the marketplace, was justified. (Pons, 1999, p. 87). Where collective sales or market-division schemes limit output, raise price, or deprive consumers of the opportunity to watch desirable games, they must be competitively justified.

American and European courts have stated that the relevant inquiry in evaluating these sort of restraints is normally the counter-factual: what would arrangements be like if competition authorities or courts did not intervene (NCAA, 1984, p. 107; Bagnasco v Banca Popolare di Novara, 1999, ¶33). Although courts have not addressed this issue, there is one area where the counter-factual approach ought not be applied. Competitive balance will rarely justify exclusively collective sales or territorial market division, because any imbalance caused by the greater marketability of individual rights sales can be redistributed through revenue sharing. It is plausible that a league may claim that, absent collective sales, transaction and bargaining costs will prevent clubs from reaching an optimal agreement on revenue sharing, and the counter-factual will, indeed, be worse than the result of collective sales. The inability of leagues to structure their affairs to reach results that efficiently maximize revenues obtainable—the inability to create a larger pie because each club cannot agree on its own fair share—should not justify anti-consumer restraints. If, indeed, this sad state of affairs persists, then perhaps further governmental intervention will be required.

Another illegitimate claim for collective sales is they will, indeed, result in greater profits from output reduction, but that the profits will be spent on worthy causes, notably grassroots development. American antitrust law has long rejected this defence. Similarly, the text of Article 81(3), requiring a ‘fair share’ of any gains from restraints to be redistributed to consumers, suggests that the ‘worthy cartel’ defence is not applicable in Europe either. Indeed, the UK recently replaced a statutory scheme that permitted a special court to find such activities to be in the public interest with one modelled on the EU Treaty. On the other hand, legislators are free to make a political choice expressly to permit worthy cartels, as was apparently done in Germany, regarding collective sales by the Bundesliga, and in France, by the statutory assignment of television rights to the Fédération Française de Football.

(iii) Labour Markets

Labour restraints imposed by leagues with monopsony power can obviously severely exploit players by limiting competitive bids to amounts significantly below what the market would bear. Three justifications are often asserted: (i) enhancing competitive balance; (ii) recovering club investment in player development; and (iii) achieving ‘cost certainty’ for clubs. Most labour restraints are unnecessary to achieve any of these goals.

Some labour restraints clearly improve competitive balance. Perhaps the best example is the North American practice of limiting the number of players each team can maintain on its roster, and requiring that players dismissed from the roster be ‘waived’ so that other teams may, if they wish, claim their contract rights. If multiple teams claim the right to the player, the team with the poorest record prevails. This rule prevents the stockpiling of players by top teams. One can imagine quite a few bottom-of-the-table Premier League teams who would insert Manchester United reserves into their starting line-up and see dramatic improvement. Another example is the NBA rookie draft. Because college, high school, and youth amateur clubs develop players,
and a single star can make a significant difference, assigning the rights to the premier amateurs based roughly on reverse order of finish in the prior season aids the poorer teams.

Salary caps and blanket restraints on signing veteran players whose contracts have expired are not tailored to promote competitive balance. Indeed, as noted above, these restraints may harm balance by inhibiting the ability of poor teams to improve quickly. This inhibition is exacerbated in North America where cash transfers are disfavoured. On the other hand, restraints on the ability of a few top teams to increase payroll might pass muster as sufficiently tailored to improving balance. This is particularly true where clubs are not profit-maximizing but are driven by members’ desire for winning at all costs.

As noted earlier, restraints have never been shown to be necessary to protect investment in players. Any club making a substantial investment in the training or development of a young player can sign the player to a multi-year contract with the last year’s terms at the option of the club. When the penultimate contract year has finished, the club can seek to re-sign a blooming player, or, if it lacks confidence it will be able to do so, the club can be compensated for transferring the player’s contract to another club willing to come to terms on a new long-term arrangement.

The justification of ‘cost certainty’, used recently in the NBA and the National Hockey League, is simply not a legitimate, pro-competitive goal for an established league. The rhetoric may have originated in 1982 when the NBA’s commissioner persuaded the players’ union to agree to a salary cap (thus exempting the cap from antitrust challenge), in part based on the need to attract additional investment into a game that was faltering. Absent viability-threatening circumstances, the desire of owners to be certain of future costs is completely antithetical to competition. Owners should be uncertain about their costs. In other industries, firms whose poor business acumen results in an inferior product do, indeed, have to spend more than their rivals to bring their goods or services up to par. In sports, when front-office management make poor personnel or coaching decisions that result in an inferior team on the field, they, too, should be required to increase spending.

The system of promotion and relegation does raise the spectre of financial insolvency if too many lower-tier teams spend money on players that only will be profitable if they secure promotion to the top tier. Maintaining the viability of lower-tier clubs might, therefore, justify rules that limit the debt that clubs can incur. However, such rules must focus on long-term debt and real threats to insolvency. A rule, for example, that limited lower-tier club payrolls to a fixed percentage of revenue prevents teams with a reasonable likelihood of promotion from spending an amount reflecting that likelihood.

V. CONCLUSION

The major sports leagues in the economically developed world rarely face competition from reasonable substitutes. Absent governmental intervention, these leagues will engage in conduct that distorts prices, reduces output, and renders output unresponsive to consumer demand. The fact that most leagues are organized as joint ventures exacerbates this problem, as inefficiency is propounded by the inability of clubs to agree on schemes that maximize profits for the entire sport because of difficulties in agreeing on how to share the increased bounty. Competition law—the English common law of restraint of trade, the American Sherman Act, and the competition articles of the EU Treaty—provides a meaningful constraint on the exercise of power by dominant sports leagues. Although the words of the statutes and judicial precedents differ, each of these systems applies the same basic test: when restraints are anticompetitive, they are illegal unless shown to be reasonably necessary to achieve a goal that is efficient or enhances consumer appeal.

This basic test should be applied with greater vigour on both sides of the Atlantic. North American consumers’ welfare would be significantly enhanced if the major sports were required to open their league to promotion and relegation, and if clubs were not allowed to maintain exclusive geographic territories for the sale of television rights. European football fans’ welfare would be significantly enhanced if collective selling were strictly limited, and
any new innovations in sporting competitions be required vigilantly to maintain meaningful promotion and relegation. Association football, rugby, and cricket may continue to be hard sells in America, and baseball and American football may show only toe-holds in Europe, but innovative systems of organizing sporting competition and effective government regulation to prevent monopolistic abuses should know no boundaries.

REFERENCES


CASES CITED

Colgate v. Bacheler (1601) 78 ER 1097.
Eastham v. Newcastle United Football Club, [1963] 3 All ER 139.
General Leaseways, Inc. v. National Truck Leasing Association, 744 F.2d 588 (7th Cir. 1984).
Greig v. Insole, [1978] 3 All ER 449 (Ch.).
McCourt v. California Sports, Inc., 600 F.2d 1193 (6th Cir. 1979).
Mitchel v. Reynolds (1711) [1558-1774] AllER 26.
Petrofina (Great Britain) Ltd v. Martin, [1966] 1 All ER. 126 (CA).
Stevenage Borough Football Club Ltd v. The Football League Ltd, unreported (Ch. 23 July 1996), transcript reprinted in The Times, 1 August 1996.
United States v. Addyston Pipe & Steel Co, 85 F.2d 271 (6th Cir. 1898).