The Sports Business as a Labor Market Laboratory

Lawrence M. Kahn

Professional sports offers a unique opportunity for labor market research. There is no research setting other than sports where we know the name, face, and life history of every production worker and supervisor in the industry. Total compensation packages and performance statistics for each individual are widely available, and we have a complete data set of worker-employer matches over the career of each production worker and supervisor in the industry. These statistics are much more detailed and accurate than typical microdata samples such as the Census or the Current Population Survey. Moreover, professional sports leagues have experienced major changes in labor market rules and structure—like the advent of new leagues or rules about free agency—creating interesting natural experiments that offer opportunities for analysis.

In this paper, I focus on four areas of research on sports labor markets. First, using data on the rise and fall of rival sports leagues and the granting of free agency rights in professional sports, economists have tested whether employer monopsony lowers worker pay. Second, sports salary and performance data, and the knowledge of the race of each player in a sport, allow us to estimate the extent of discrimination in a much more detailed way than is possible in other industries. Third, changes in rules about the draft and player movement allow us to examine the claim of the Coase theorem that with complete information but no transactions costs or wealth effects, the ownership of a resource has no impact on the allocation of resources; rather, changes in such laws will only affect the distribution of wealth (Posner and Parisi, 1997). Finally, because of the observability of contract provisions, supervisor quality and worker behavior, one can estimate the impacts of supervision and incentives on behavior.

Of course, it is wise to be hesitant before generalizing from the results of sports research to the population as a whole. The four major team sports employ a total

Lawrence M. Kahn is Professor of Labor Economics and Collective Bargaining, School of Industrial and Labor Relations, Cornell University, Ithaca, New York.
of 3,000 to 4,000 athletes who, in the mid-1990s, earned an average of $600,000 in hockey, $650,000 in football, $1.2 million in baseball, and $1.8 million in basketball (Staudohar, 1996), far above the 1997 median earnings of full-time, full-year equivalent workers of $25,000–$26,000 (U.S. Bureau of Labor Statistics, 1998). But at a minimum, sports labor markets can be seen as a laboratory for observing whether economic propositions at least have a chance of being true. Moreover, I have found sports to be an excellent tool to teach students about economics and get them interested in the subject. Concepts like the marginal revenue product come alive when students are talking about their favorite players.

Monopsony and Player Salaries

Sports owners are a small and interconnected group, which suggests that they have some ability to band together and act as monopsonists in paying players. The result is that player pay is held below marginal revenue product. I discuss three sources of evidence on monopsony in sports: evidence from the rise and fall of rival leagues, evidence from changes in rules about player free agency, and studies comparing the marginal revenue product of players with their pay. Sports owners have often had monopsony power over players in the sense that in many instances players have the option of negotiating only with one team. Here, salaries are determined by individual team-player bargaining in which marginal revenue product, and the outside options available to teams and players, will affect the outcome. Rules changes and the rise and fall of rival leagues have their effects by changing players’ and teams’ outside options.

Rival Leagues

There have been two time periods in which rival leagues posed a substantial threat to existing professional sports. The first is the period from 1876 to 1920, when there was a scramble of professional baseball leagues forming, merging, and dissolving. The second is the period from the late 1960s into the early 1980s, when new leagues were born in basketball, hockey, and football.

Figure 1 shows the pattern of salaries for major league baseball players during the period from 1876–1920. Baseball is the oldest major league sport in the United States, beginning with the birth of the National League in 1876.¹ In this early period, there was competition for player services from other baseball leagues. To protect itself against the competition of rival leagues and improve the team owners’ balance sheets, the National League introduced the “reserve clause” in 1879, which meant that players were bound to the team that originally acquired the rights to contract with them. Owners now had additional monopsony power over players, and player salaries dropped.

However, the lower salaries may have contributed to the birth of a new league in 1882, the American Association. Figure 1 shows a rapid escalation of player salaries during the 1882–91 period, one that was characterized by many instances of player

¹ This discussion of early baseball history is based on Burk (1994) and Staudohar (1996).
movement across the two competing leagues. Average nominal National League salaries rose from $1,375 in 1882 to $3,500 in 1891, which would be equivalent to about $63,000 in 1998 dollars (U.S. Bureau of the Census, 1975; BLS website).²

This increase in salaries is not conclusive evidence that monopsony power of owners decreased; after all, salaries could have risen for other reasons, like the growth of baseball’s popularity. However, in 1891, four of the American Association teams were absorbed into the National League, and five dissolved AA franchises were bought out by the survivors. Figure 1 shows an abrupt, massive decline in National League player salaries starting in the first season of the merger: player pay fell from $3,500 in 1891 (before the merger) to $2,400 in 1892 to $1,800 in 1893. This pay cut was accomplished as the outcome of the National League owners’ announcement in 1893 of a new salary policy: the maximum pay for a player was to be $2,400. Indeed, some teams imposed lower caps: eight top players on the Philadelphia Phillies who were all paid more than $3,000 in 1892 found that they were all paid exactly $1,800 in 1893 (Burk, 1994, p. 125). The sharp decline in player salaries does not appear to reflect a major decline in the demand for baseball entertainment, as attendance climbed through the 1895 season.

The success of the National League waned somewhat in the late 1890s, partly due to a lack of competitive balance, but the baseball market was growing. A new rival league, the American League, began in 1901 with eight teams. It successfully lured many star players from the older league and actually outdrew it in attendance in 1902 by 2.2 million to 1.7 million (Burk, 1994, p. 151). In response, the National League attempted to have its reserve clause enforced by state courts to prevent players from jumping leagues; however, because state courts have no jurisdiction for player movements outside a given state, it was ultimately unsuccessful in this effort (Burk, 1994; Berry, Gould and Staudohar, 1986).

² These salary figures come from unpublished data used in Burk (1994, Appendix Figure 1) courtesy of Robert Burk.
A familiar pattern then emerges. The huge success of the American League brought with it a dramatic rise in player salaries. In fact the salary increase appears to begin in 1900, perhaps reflecting anticipation of the new league. The two leagues merged during the 1903 season, at the end of which the first World Series was played. Then in 1903, salaries in Major League Baseball fell immediately by about 15 percent. (In Figure 1, the data from 1903 and beyond refer to American and National League teams combined, while the pre-1903 data are for the National League only.) Again, this decline does not seem to reflect any fall in baseball’s popularity that year.

Major League Baseball prospered for the rest of the first decade of the twentieth century, with player costs under control and attendance on the rise. However, attendance fell beginning in 1910, and owners kept a tight lid on salaries. Player discontent resulted in the formation of a union, the Fraternity of Professional Baseball Players of America, at the end of the 1912 season. The owners were under no legal obligation to bargain with a union and reacted to it with some hostility.

This dissatisfaction among players helped pave the way for the Federal League, which was able to recruit to long-term contracts several well-known Major League ballplayers beginning in 1913. The pay cycle began again. While the Federal League was in existence from 1913 through 1915, many players jumped leagues, and major league salaries went from about $3,000 in 1913 to $5,000 in 1915. After the 1915 season, most of the Federal League’s owners were “bought out” in December 1915 by the major leagues, and nominal salaries plummeted back to $4,000 by 1917, a fall that was even larger in real terms due to the inflation of the World War I period.

The Federal League owners who were not part of the settlement pursued an antitrust suit against the settling parties for creating a monopoly; however, this suit was lost in 1922, when the U.S. Supreme Court declared that baseball was not a business (Federal Baseball Club of Baltimore, Inc. vs. National League of Baseball Clubs, et al., 259 U.S. 200 [1922]). This decision began baseball’s antitrust exemption, which was upheld several times, most notably in an unsuccessful attempt by a player named Curt Flood to become a free agent in 1969 (Flood vs. Kuhn, 407 U.S. 258 [1972]). While Flood lost his case, it may have set the stage for baseball players’ ultimately successful quest for free agency. However, this goal was not achieved through the antitrust laws; in fact, baseball’s exemption lasted with respect to player relations until legislation ending it was passed in 1998. Rather, collective bargaining brought free agency to baseball, as discussed below.

The early experiences of Major League Baseball provide some compelling evidence for the potential impact of monopsony in this labor market. However, the comparisons just discussed all concern baseball, and thus have no real control group. In the modern period, we can use salaries in some sports as control groups for other sports.

From the late 1960s into the 1970s, highly credible rival leagues were born in basketball and hockey. The American Basketball Association (ABA), which lasted from 1967 to 1976, was able to field some very good teams. In 1976, four of its teams were absorbed into the NBA, and these all made the NBA playoffs in several seasons after the merger. The NBA Players Association (NBAPA) challenged the merger on antitrust grounds, but then withdrew its lawsuit as the result of a settlement which granted free agency rights to NBA players. The World Hockey Association (WHA),
which lasted from 1971 to 1979, also had several excellent teams which were absorbed into the NHL starting in 1979.

The rise and fall of these two rival leagues offer another opportunity to test how monopsony might affect salaries since the other two major team sports—baseball and football—had no such competition in their labor markets until the advent of free agency in baseball in 1976 and the birth of the United States Football League (USFL) in 1982.

Table 1 shows average player salaries for the major team sports for selected years from 1967 to 1982. In 1967, there were no rival leagues in baseball, football, or hockey, and the ABA was just getting started. Further, there was no free agency, and players’ unions had not yet negotiated their first agreements. Thus, the 1967 salaries can be viewed as representing common initial conditions with respect to negotiating rules, although not necessarily with respect to demand conditions.3

By 1970 and 1972, the ABA had been in existence for several years, and NBA players rapidly became the highest paid of the major team sports. The World Hockey Association started in 1971, and by 1972, NHL players outearned football and baseball players by similar margins. These upward movements in the relative salaries of basketball and hockey players, while consistent with effects of the new competition, need to be judged against the changing popularity of these two sports. For example, in the NBA, total attendance rose by 120 percent between 1966–67 and 1971–72, while television revenues went from $1.5 million to $5.5 million during the same period, rises that were much, much greater than the increases for football or baseball during this time (Quirk and Fort, 1997). The shock of higher salaries may also have spurred the teams to market themselves better. Moreover, NBA salaries as a percentage of gross basketball revenues rose from 30 percent in 1967 to 66 percent in 1972, which suggests a structural shift in salary determination that goes beyond a rise in revenues (Staudohar, 1996, p. 108). In the NHL, attendance growth was actually much faster in the five years before the birth of the WHA (135 percent) than while the WHA was in business (7 percent) (Quirk and Fort, 1997). Thus, the acceleration of NHL salaries after 1971 is telling indeed.

There is one more important example of the impact of a rival league on player salaries, the United States Football League (USFL). Like some of the baseball experiences earlier in the century, the USFL was born out of labor strife in the established league; in this case, it was a seven-week NFL strike in 1982, in which the players had failed to gain any significant ground in their fight for free agency or a share of revenues. From 1982 to 1985, the USFL posed a challenge to the established NFL. The USFL had strong financial backing from such owners as Donald Trump; many NFL players switched leagues; and the USFL was able to sign some high-profile college players such as Anthony Carter and Herschel Walker (Staudohar, 1996, p. 64). However, poor television ratings for the USFL ultimately signaled its demise.

The pattern of football player salaries during the USFL years follows that set by

---

3 The higher NFL salary level may have represented the lingering effects of the bidding war between the National Football League and the American Football League earlier in the 1960s. The two leagues agreed to merge in 1966, a merger that required congressional approval (Ross, 1989). This episode is not discussed here because NFL salary level information is not available before 1967. However, there is strong anecdotal evidence of bidding wars between the two leagues over college stars such as Joe Namath (Neft and Cohen, 1988, p. 94).
the example of other rival leagues. Real salary increases for NFL players averaged 4 percent per year from 1977–1982, before the USFL, and 5 percent per year from 1985–1989, just after the USFL. In between, real player salary increases were 20 percent per year from 1982–1985 (Flanagan, 1989; Staudohar, 1996). Changes in the popularity of football do not seem sufficient to explain the explosion of salaries during the USFL years. From 1977 to 1981, NFL attendance rose 23 percent, but attendance was actually 2 percent lower in 1985 than in 1981; further, NFL television revenues rose by similar rates before and during the USFL years. Overall, NFL salaries during the USFL period were much higher than could have been predicted on the basis of revenues during that time (Flanagan, 1989).

If one uses other sports as a control group for the experience of football salaries in the USFL years, the same lesson emerges. The salary growth for football players from 1982 to 1985 was 8–10 percentage points per year higher than baseball and basketball, and 17 percentage points per year higher than for hockey players (Flanagan, 1989; Staudohar, 1996). However, during the 1981–85 period, attendance grew 5 percent in baseball, 11 percent in the NBA, 11 percent in the NHL, and, as noted, fell 2 percent in the NFL; television revenues grew 211 percent in baseball and 35 percent in the NBA, compared to 29 percent in the NFL (Quirk and Fort, 1997). The faster growth of NFL salaries during the 1982–85 period despite worse attendance and television revenue increases again suggests the importance of the USFL.

**Free Agency**

Until 1976, players in each of the four major sports were bound by the reserve clause to remain with their original team, unless that team decided to trade or sell them to another team. They were not allowed to become free agents, who could sell their services to any team.

The path toward free agency started in baseball. The Major League Baseball Players Association (MLBPA) was started in 1952, but it didn’t become a modern union

---

4 Deceleration of the growth of NFL television revenues from nearly 30 percent per year from 1981 to 1985 to just 2.6 percent annual growth from 1985 to 1988 could also have explained the deceleration of salaries after 1985. However, it is also true that attendance in the NFL picked up from 1985 to 1988, rising 1.5 percent, in contrast to the 2 percent fall from 1981 to 1985 (Quirk and Fort, 1997).
until the former United Steelworkers negotiator, Marvin Miller, took over in 1966.\textsuperscript{5} The MLBPA achieved a collective bargaining agreement in 1968, and in 1970 the National Labor Relations Board ordered the parties to use outside arbitrators for resolving grievances. In a farsighted decision, Miller obtained management’s agreement to incorporate the standard player contract, which included the reserve clause, into the collective bargaining agreement. This meant that grievances about the interpretation of this clause were a proper subject for arbitration. In December 1975, an arbitrator ruled that the reserve clause meant that the team could reserve a player for only one year beyond the expiration of any current contract. With the reserve clause in place, almost all teams signed players exclusively to one-year contracts, and so this ruling would have freed virtually all of the veteran players after the 1976 season. The teams, threatened with this possibility, were thus moved to negotiate a formal system of free agency with the union in 1976, calling for free agency (with some relatively minor compensation to any team losing a free agent) for players with at least six years of Major League Baseball service. This provision remains basically in place. In 1994, about 33 percent of players had at least six years’ service (Vrooman, 1996, p. 354).

The rise of free agency in the 1976–77 period had a powerful impact on the salaries of baseball players. The average real increase in baseball salaries was from 0–2 percent per year from 1973–75. In 1976 the average real salary increase was almost 10 percent; in 1977, the first year under the new collective bargaining agreement, 38 percent (!); in 1978, 22 percent, before falling back into single digits growth in 1979 (Lehn, 1990). Moreover, baseball salaries as a percentage of team revenues rose from 17.6 percent in 1974 to 20.5 percent in 1977 to 41.1 percent in 1982, further suggesting that free agency has had a structural effect on baseball salary determination (Zimbalist, 1992).\textsuperscript{6}

A final point to note about baseball salary determination is that in a series of grievance arbitration decisions in the 1980s, the owners were found guilty of colluding by not making offers to free agents (Staudohar, 1996). This reassertion of cartel wage-setting behavior appeared to be successful in restraining salary growth. Real annual growth in baseball salaries fell from 11 percent in the 1982–85 time period to 3 percent from 1985–87 (Flanagan, 1989; Staudohar, 1996; BLS website). Moreover, salaries as a percent of revenues fell from about 40 percent in 1985 to 32 percent in 1989 during the collusion period (Zimbalist, 1992). In 1989, arbitrators levied a $280 million back pay penalty on the owners to be paid out over the 1989–91 period as compensation for the losses imposed by collusion (Staudohar, 1996), and salaries as a percent of revenue bounced back to 43 percent by 1991 (Zimbalist, 1992). The collusion episode provides a further illustration of the potential impact of monopsony on salaries.

\textsuperscript{5} This description of the Major League Baseball Players Association and baseball labor relations is based on Staudohar (1996), Berry, Gould and Staudohar (1986), and Korr (1991).

\textsuperscript{6} In the 1974 and 1975 seasons, many players were eligible for salary arbitration, but without free agency, arbitration alone seemed to have little impact on average salaries. However, salary arbitration, in combination with free agency, can have an effect on players’ average salaries, because arbitrators are often entitled to compare players’ demands and team offers to the salary levels of free agents. Using fixed effects methods on panel data for individual players for the late 1980s, I found that, other things equal, being eligible for salary arbitration increased players’ salary by about 30–45 percent. See Kahn (1993a).
Basketball players also won free agency in 1976, but by a different route, through the settlement of the players’ antitrust suit challenging the ABA-NBA merger. As a result, free agency in the NBA came on the heels of the ABA-NBA salary war period of 1967–76. Possibly as a result, average salaries grew more slowly in the NBA in the 1977–82 period than in football or baseball, a comparison which does not suggest a major impact of free agency in the NBA in addition to the impact of competition from the ABA. On the other hand, NBA salaries amounted to about 70 percent of revenue in 1977 and “nearly three quarters” in 1983, suggesting some further increase in basketball players’ relative power after the coming of free agency even without the benefit of an alternative league (Staudohar, 1996, p. 108). Finally, basketball imposed a salary cap in 1983, and salaries did indeed decelerate after 1985. However, there were many exceptions to the salary cap, and it may ultimately have had little effect on salaries during this period (Fort and Quirk, 1995).

**Evidence on the Degree of Monopsonistic Exploitation**

To this point, the argument has relied on presenting abrupt shifts in salaries that are difficult to explain without appealing to the theory of monopsony. An alternative mode of research on salary determination is to compare estimates of players’ marginal revenue products to salaries, and in this way to approximate the degree of monopsonistic exploitation. This research methodology was pioneered by Scully (1974) in work on baseball and has been elaborated in many papers, summarized in Scully (1989) and Zimbalist (1992). Scully’s (1974) contribution also has inspired a large number of undergraduate economics theses.

Scully’s (1974) approach involved a two-step process. First, one estimates how various measures of a player’s performance (like runs scored and batted in, or ratio of batters struck out to batters walked) affect a team’s winning percentage. Second, one estimates how team revenue is affected by winning percentage, along with other factors like the size of the local market. A player’s marginal revenue product is estimated by multiplying his contribution to winning by the impact of winning on revenue. Using this methodology for players in the reserve clause days of the late 1960s, Scully (1974) estimated that average players received about 20 percent of their marginal revenue product and stars only about 15 percent. Scully (1989) and Zimbalist (1992) updated and refined Scully’s (1974) calculations for the late 1980s. Scully (1989) found that star players in 1987 were paid 29–45 percent of marginal revenue product; even though this was the height of the collusion period, the percentage was still much higher than the 15 percent he found for the reserve clause days. Zimbalist’s (1992) approach compares the players eligible and not eligible for free agency. In 1989, for those players with less than three years’ service, and thus not eligible for salary arbitration or free agency, the ratio of salary to marginal revenue product was just .38 times what it was for those eligible for salary arbitration only and .18 times that for those eligible for free agency.7

These measures of monopsonistic exploitation must be interpreted cautiously

---

7 The low relative levels of exploitation for the free agency eligible suggest that in the free agent market, teams may have been affected by the “winner’s curse.” For more evidence on the winner’s curse in the baseball market for free agents, see Cassing and Douglas (1980).
since (as noted by the authors of the studies) they do not control for a player’s effects on revenue other than through his own playing statistics’ effects on winning.\footnote{As an extreme example of such effects not captured by the revenues of the player’s home team, Hausman and Leonard (1997) estimate Michael Jordan’s value to other NBA teams during the 1991–92 season to be roughly $53 million. This consisted of effects on attendance at away games, television ratings, and merchandise sales.} However, taken as a whole, this line of research produces additional evidence that making the labor market more competitive leads to higher salaries than would be the case under monopsony. Nonetheless, during the 1980s there still appeared to be widespread monopsonistic exploitation in baseball, and research from this period also showed similar results for basketball (Scott, Long and Somppi, 1985).

### Racial Discrimination in Professional Sports

Gary Becker (1957), in his seminal work on the economics of discrimination, identified three forms of discrimination: employer, coworker (including supervisors and supervisees as well as lateral coworkers), and customer. Under certain conditions, discriminating employers will be driven out of business by the nondiscriminator(s), and coworker discrimination will lead to equally competitive segregated firms with equal pay for equal work. However, customer discrimination is different from these two forms of prejudice, because an employer that pays more money to the types of workers whom customers prefer is likely to be rewarded by the market. Conversely, those workers who are not preferred by customers, but have a personal comparative advantage in the affected job, will either need to accept lower pay, if they remain in the customer sector, or else move to the noncustomer sector, where their comparative advantage is lost. This argument suggests that competitive forces are less likely to eliminate customer-based discrimination than that based on employer or coworker prejudice. With heterogeneous consumer preferences, we may observe segregation and equal wages, if consumers have equal income and access to capital. Thus, the persistence of wage differences based on customer discrimination will be an empirical question.

Of course, the sports industry is a customer-based service sector. Further, sports leagues as monopolies may not face the kind of free entry that might serve to discipline discriminating employers. Until the 1940s, black players were banned from major league sports in the United States.\footnote{In the 1880s, several black players were allowed into the American Association, which vied with the National League until its demise in 1891. But the National League, and later the American League, enforced a color bar until Jackie Robinson’s 1947 rookie season with the Brooklyn Dodgers (Burk, 1994; Tygiel, 1983).} Baseball began slowly integrating with the hiring of Jackie Robinson by the Brooklyn Dodgers in 1947, and football and basketball soon followed suit. By the mid-1990s, the population of major league baseball players was about 30 percent black, the NFL was 65 percent black, and black players were 80 percent of the NBA’s ranks (Staudohar, 1996). But although the color barrier has been decisively broken, it is still possible that racial discrimination could persist on the margin.
If most employers have a taste for discrimination, then a nondiscriminating employer can either win more games, or reduce expenses, or both, by hiring the most productive workers at the lowest cost possible. In the context of Major League Baseball’s integration, some teams were more active in bringing black players onto their rosters in the late 1940s and 1950s (Gwartney and Haworth, 1974; Hanssen, 1998). Those teams that were more willing to use black players did indeed have significantly more success on the playing field than those that did not.

The labor market manifestations of discrimination can include salary discrimination (unequal pay for equal work), hiring discrimination, retention discrimination, and positional segregation. A considerable volume of research has attempted to determine the extent of these outcomes; Kahn (1991) offers a survey of this evidence through the 1980s.

Among the forms of discrimination in sports, salary discrimination is the most studied issue. The typical research design—similar to much work in this area in labor economics—is a regression in which log salary is the dependent variable, and the independent variables include performance indicators, team characteristics, and market characteristics, with a dummy variable for white race. If the coefficient on the white indicator is positive and significant, then this potentially offers evidence of discrimination. Alternatively, some researchers have used separate regressions for white and nonwhite players, testing the possibility that performance is rewarded differently by race.

A major difficulty for all labor market research on discrimination is the problem of unobserved or mismeasured variables, such as the quality of schooling among workers in general. However, such problems must surely be less severe in sports than elsewhere. For example, the Baseball Encyclopedia and other baseball data sources allow one to control for very detailed performance indicators like batting average, stolen bases, home runs, career length, team success, and many more. “Occupation” in baseball is one’s position, a far more detailed indicator than, say, “machine operative.” The accuracy of the compensation data in sports, in many cases supplied by the relevant players’ union that keeps copies of the actual player contracts, is very high.

The sport where regression analyses have produced the most evidence of salary discrimination is professional basketball. In the mid-1980s, several studies found statistically significant black salary shortfalls of 11–25 percent after controlling for a variety of performance and market-related statistics (for example, Kahn and Sherer, 1988; Koch and Vander Hill, 1988; Wallace, 1988; Brown, Spiro, and Keenan, 1991). However, by the mid-1990s, there were no longer any overall significant racial salary differentials in the NBA, holding performance constant (Hamilton, 1997; Dey, 1997; Bodvarsson and Brastow, 1998). One caveat to this finding is seen in Hamilton’s (1997) results from quantile regressions, which estimate the extent of discrimination at dif-

---

10 While there may still be omitted variables that could have explained the *ceteris paribus* white salary advantage, reverse regression tests, which can under some restricted circumstances take account of such problems (Goldberger, 1984), showed even larger apparent discrimination coefficients against black players (Kahn and Sherer, 1988). In fact, these larger effects suggest that black players had better unmeasured productivity characteristics than whites, at least under the statistical assumptions outlined by Goldberger (1984).
ferent points of the salary distribution, conditional on productivity. He did not find evidence of discrimination at the 10th, 25th, and 50th salary percentiles, but there was a significant white salary premium of about 20 percent, other things equal, at the 75th percentile of the salary distribution and above.\footnote{11}

Customer preferences may have something to do with the racial pay gap observed in basketball in the 1980s. For example, Kahn and Sherer (1988) found that, all else equal, during the 1980–86 period each white player generated 5,700 to 13,000 additional fans per year. The dollar value of this extra attendance more than made up for the white salary premium, a finding consistent with the existence of monopsony. Other researchers found a close match between the racial makeup of NBA teams in the 1980s and of the areas where they were located, again suggesting the importance of customer preferences (Brown, Spiro and Keenan, 1991; Burdekin and Idson, 1991; Hoang and Rascher, 1999). However, by the 1990s, customer preferences for white players were less evident. Dey (1997), for example, found that all else equal, white players added a statistically insignificant 60 fans apiece per season during the 1987–93 period. This evidence is consistent with the decline in the NBA’s overall unexplained white salary premium from the 1980s to the 1990s, although Hamilton’s (1997) results suggest that it is possible that white stars add fans even if the average white player does not.

If NBA fans do have preferences for white players, having white benchwarmers may be a cheap way for teams to satisfy such demands. While early research found that white benchwarmers had longer careers than black benchwarmers (Johnson and Marple, 1973), more recent work does not find that benchwarmers are disproportionately white (Scott, Long and Somppi, 1985).

In contrast to these findings in basketball, similar regression analyses of salaries in baseball and football have not found much evidence of racial salary discrimination against minorities. For example, in baseball, these kinds of analyses never seem to find a significantly positive salary premium for white players. Among nonpitchers, some studies actually have found significantly negative effects of being white in the late 1970s and 1980s (Christiano, 1986, 1988; Irani, 1996); however, my own reanalysis of the same data used in one of these studies found that these differentials disappeared when a longer list of productivity variables was added (Kahn, 1993a). In football in 1989, Kahn (1992) found only very small salary premia (discrimination coefficients) in favor of whites of only 1–4 percent, and these differences were usually not statistically significant. However, nonwhite NFL players earned more in areas with a larger relative nonwhite population than nonwhites elsewhere, and whites earned more in more white metropolitan areas than whites elsewhere. These findings suggest the influence of customers, but they did not add up to large overall racial salary differences in the NFL.\footnote{12}

\footnote{11} Logically, if there is no \textit{ceteris paribus} pay gap on average, and white stars receive a premium, then blacks at the bottom should earn more than whites. Hamilton (1997) finds point estimates in this direction, but they are not statistically significant.

\footnote{12} A lively literature has developed on the issue of discrimination against French Canadians in the National Hockey League. Some authors have found apparent salary discrimination against this group in
Although little evidence exists of a discriminatory salary premium in baseball or football, there is evidence of other forms of discrimination in sports, some of it among customers, and also in hiring, retention and assignment of players. For example, a larger number of white players seems to lead to added baseball fans, over the time period from the mid-1950s through the 1980s (Hanssen, 1998; Irani, 1996). Moreover, baseball cards for white players sold at a significantly higher price than those of comparable black players in 1989 (Nardinelli and Simon, 1990).

On hiring, there is indirect evidence that black players went later in the NFL draft than whites of equal playing ability during the 1986–91 period (Conlin and Emerson, 1998). However, in basketball, a study found only small, insignificant racial differences in draft order among NBA players on rosters in 1985, conditional on college performance—and these differences favored black players (Kahn and Sherer, 1988). All studies of drafts suffer somewhat from the fact that there is no systematic information available on those not drafted and not on rosters.

On player retention, Jiobu (1988) found that from 1971 to 1985, black players in Major League Baseball had a significantly higher exit rate than whites, other things equal, and Hoang and Rascher (1999) obtained a similar result for the NBA for 1980–91. A reasonable interpretation of these differences in exit rates is that they reflect team decisions not to offer players a new contract. Whether this disparity held up in the 1990s or whether the effects are similar for benchwarmers and more regular players are interesting questions for future researchers.

On positional segregation, Christiano (1988) found that black baseball players in 1987 were overrepresented in the outfield and underrepresented at infield and catcher. Kahn (1992) found that in the NFL in 1989, black players were overrepresented at running back, wide receiver and defensive back, and underrepresented at quarterback, kicker, punter and offensive line. Whether this segregation represents current discrimination, past discrimination by high school or college coaches, performance differentials or self-selection is an open question, and there is a lively if still inconclusive debate about these issues (Kahn, 1991). Whatever the cause of segregation, it can have long-term consequences. In baseball, for example, managers are drawn from the middle infield positions, which have been disproportionately white (Scully, 1989).

The Coase Theorem and Sports

In the frictionless, competitive world of the Coase theorem, we expect changes in the draft or in free agency to have no effect on player allocation (Rottenberg, 1956). To understand why, compare player allocation under the reserve clause with that under free agency. Suppose there is a player whose marginal revenue product is $500,000 in Kansas City and $1,000,000 in New York. Under the reserve clause, we

---

Canadian cities outside Quebec province, a pattern consistent with the notion of customer discrimination (Jones and Walsh, 1988; Longley, 1995). Yet others have disputed this interpretation and the findings as well (Krashinsky and Krashinsky, 1997). There is also a debate over whether French Canadians face entry barriers into the NHL (Walsh, 1992; Lavoie, Grenier and Coulombe, 1992).
expect the two teams to contract with each other to move the player to New York: any offer of cash or players in trade by New York to Kansas City that has value greater than $500,000 to the latter but less than $1,000,000 to New York will make both teams better off financially. If the player is a free agent, then New York will make a higher bid than Kansas City, and we thus expect the player to move voluntarily to New York. Player allocation is thus the same under the reserve clause and free agency; the only difference is who gets the money. There is also the possibility that owners are sports enthusiasts rather than profit maximizers (Vrooman, 1997). But in this case, a wealthy owner will overpay other teams to get good players under the reserve clause or will overpay players directly under free agency. Again, there is no reason here for player allocation to be any different under free agency and the reserve clause.

This reasoning has powerful implications for sports, since teams have often claimed that free agency needs to be limited, or else the rich teams will buy all of the good players and competitive balance will supposedly be destroyed. The Coase theorem suggests that free agency will not affect player allocation, but this result is based on the assumptions that there is perfect information, transactions costs are zero and that wealth effects are zero. If these assumptions do not hold, then rules changes can have real effects (Daly and Moore, 1981; Lehn, 1990).

The transactions costs argument suggests why sports leagues may impose barriers to player movement. For example, in 1976, the Commissioner of Major League Baseball, Bowie Kuhn, voided the sale of three star players (Vida Blue, Joe Rudi and Rollie Fingers) from the Oakland Athletics to the (already) league champion Boston Red Sox, citing an unreasonable effect on competitive balance (Quirk and Fort, 1997). If trades are more difficult to arrange than cash sales, then the Commissioner’s discouragement of these sales may have affected competitive balance. Also, if trades are more difficult to arrange than cash sales, the allocation of players may be different under free agency, where cash is the mechanism for moving players, from that prevailing under the reserve clause. Overall team salary caps imposed by the league can be seen as a way of building in transactions costs by making it more difficult for teams with high payrolls to acquire even more expensive players.

If wealth effects are important, they can produce a different allocation of resources under different rules for assigning property rights in a sport (Lehn, 1990). As we have seen, granting players free agency in baseball has raised player wealth enormously. With this increased wealth, player willingness to pay to be in favorable locations or on good teams may be higher than it would have been under the reserve clause with its lower salaries. Therefore under free agency, teams in “favorable” locations—however this is defined by players—have a larger comparative advantage over teams in less preferred areas than under the reserve clause. Ultimately, competitive balance need not be affected; if more teams are located in these favored areas, then the player talent can then be spread evenly across these teams. On the other hand, if willingness to pay to be on a good team rises with wealth, then competitive imbalances may last longer under free agency.

13 The tax code also discourages cash sales, since a cash sale may result in a capital gains tax for the selling team; however, trades have no associated capital gains for tax purposes (Quirk and Fort, 1997).
Research on the Coase theorem in sports seeks to determine whether changes in rules have had real resource allocation effects. If such effects are found, then it is plausible to argue that certain assumptions of the Coase theorem do not hold. Two types of evidence are typically sought in these studies: the impact of rules changes on competitive balance, or the effect of rules changes on the mobility of individual players.

Research on the impact of the draft or free agency has—with two notable exceptions—generally concluded that these rules do not affect competitive balance. For example, the National Football League in 1936 instituted a player draft as the result of a bidding war over a college star at the end of the 1934 season, with the worst teams drafting first. Fort and Quirk (1995) in fact find that the draft had no effect on the standard deviation of winning percentage in the NFL during this period.

A similar analysis of competitive balance in Major League Baseball before and after the introduction of free agency indicates no significant changes in the standard deviation of winning percentages, in the 1966–75 vs. 1976–85 periods (Fort and Quirk, 1995). In basketball, the gap between the winning percentage of the top team and the bottom team in the National Basketball Association was virtually identical during the reserve clause-pre-ABA days of 1961–67, the ABA period of 1967–76, and the free agency period of 1976–83 (Noll, 1991). These findings do not suggest any impact of free agency (or the rival ABA) on competitive balance. These findings are consistent with the Coase theorem in that rules changes are not associated with changes in competitive balance. Of course, if teams are unable to judge talent, then changing the rules will not have any effect on competitive balance. But the systematically better performance of particular managers suggests that some teams can in fact do this (Kahn, 1993b).

There are two notable exceptions to these findings that changes in rules have had little effect on competitive balance, both relating to baseball. Both exceptions may be susceptible to alternative explanations; still, they are worth noting. First, after the introduction of the baseball draft in 1965, the competitive balance in the American League appears to have increased markedly. Fort and Quirk (1995) find in a study of baseball from 1952 to 1975 that the standard deviation of winning percentage was significantly lower in the American League in nonexpansion years after the draft than before the draft. The concentration of championships fell in both the National and American Leagues after the introduction of the draft, suggesting that it in fact did affect competition. But it is also true that from 1952 through 1964, the 13 seasons before the draft, the Yankees won 11 pennants—and then did not win any pennants until 1976. Thus, the change in the American League may reflect the downfall of the Yankees more than the draft; indeed, it has been argued that the Yankees’ apparent reluctance to employ black players in the 1960s contributed to their decline (Hanssen, 1998). In the National League, there was hardly any change in the standard deviation of winning percentage.

A second piece of evidence that rules changes may have real effects suggests that free agency in baseball has lowered the year-to-year correlation of winning percentage by breaking up winning teams (Vrooman, 1996). For example, last year’s winning percentage had an effect on this year’s winning percentage of .715
from 1970–76 before free agency; the coefficient was then .711 from 1978–80, .439 from 1983–85, .260 from 1986–89, and .036 from 1990–93, with all effects significant except the last one. This evidence is intriguing, but uncertain. While the year-to-year correlation of winning percentage has fallen, there is no sharp decline around 1976, and then when the decline does occur, it continues into the early 1990s. The timing suggests that the introduction of free agency in 1976 may not have been the only factor at work here. Other causes might have included changes in the revenue position of teams, which would by themselves affect competitive balance under the reserve clause or free agency.

A more direct test of the Coase theorem than looking at competitive balance is to examine player mobility under different rules for player allocation. Hylan, Lage and Treglia (1996) examined the mobility of all major league baseball pitchers over the 1961–1992 period, so that the beginning of free agency in 1976–77 is at the midpoint of this interval. Before 1976, total player service time was significantly positively associated with mobility, other things equal; after 1976, this was still true, but the effect of longer service on mobility was much smaller than before free agency. This outcome could have occurred if free agency has led to a greater incidence of long-term contracts for more experienced players, as suggested by Lehn (1990) and Kahn (1993a). Wealth effects, due to free agency, could have raised the willingness of players to pay for amenities by staying put.

The mobility evidence appears to contradict the stereotype that while in the old days (before free agency), players stayed with one team, but now players are mercenaries and will move at the drop of $50 million. However, we often forget that there were many trades before free agency, and once a player moves nowadays to a team offering him a long-term contract—the mercenary stereotype—that player then typically becomes relatively immobile.

**Incentives, Supervision and Performance**

Some of the most intriguing evidence on the links from incentives to performance comes from sports that have not been much discussed to this point, like golf and marathon running. Ehrenberg and Bognanno (1990a, b) used data from the 1984 U.S.-based Professional Golf Association (PGA) and 1987 European PGA tours to estimate the impact of incentives on player performance. Because the prize structure of a given tournament is known in advance, one can compute the dollar gain to improving one’s finishing position in a tournament. Ehrenberg and Bognanno found that a greater dollar gain to a better finish had a statistically significant favorable effect on a player’s performance, controlling for the player’s ability, his opponents’ ability, and the difficulty of the course. In addition, golfers appear to perform better when it matters more, particularly in the later rounds of a tournament. Finally, golfers’ labor supply, as measured by their propensity to enter a given tournament, is positively affected by the expected gain to participating, implying an upward-sloping labor supply
However, a more recent replication study, using 1992 PGA data, found that monetary incentives had small and statistically insignificant effects on player performance and that results were sensitive with respect to who rated the weather that prevailed during a tournament (Orszag, 1994).

The framework devised by Ehrenberg and Bognanno (1990a, b) has been used to examine the incentive impact of prize money in two additional sports: marathon running and auto racing. In auto racing, Becker and Huselid (1992) found that larger monetary rewards to better finishes lowered individual racers’ finishing times and raised the incidence of accidents, presumably due to a greater effort to go fast. In marathon races, Frick (1998) found that better prize money and performance bonuses for setting records lowered racing times.

In the major team sports that have been the primary focus of this paper, free agency has brought with it an increased incidence of long-term contracts, a finding Lehn (1990) argued was consistent with wealth effects, as players in essence buy long-term income insurance. He noted that as the incidence of long-term contracts went from virtually zero during the days of the reserve clause to 42 percent of baseball players with at least two years pay guaranteed as of 1980, the share of baseball players who spent time on the disabled list rose from an average of 14.8 percent from 1974 to 1976 (before free agency) up to 20.8 percent from 1977 to 1980 (the early years of free agency). Lehn surmised that this increase was a moral hazard response by players on guaranteed long-term contract. In this instance, moral hazard refers to a player’s impact on the decision to go or stay on injured reserve.

To perform a sharper test of this hypothesis, he compared players in 1980 who had long-term contracts of three years or more with those who had short-term contracts of two years or less. Prior to signing these contracts, those with long-term contracts were almost two years younger and had 2.2 days per season less disability than those who signed short-term contracts. Thus, those with long-term contracts do not appear to be an especially injury-prone group. Nonetheless, after signing their agreements, those with long-term contracts averaged 12.6 disabled days per season, compared to only 5.2 days for those with 0–2 years. Lehn (1990) confirms in a regression setting that this effect is highly statistically significant. The finding is strongly suggestive of a moral hazard effect, although one cannot completely rule out that players who had private information that they were fragile were more likely to sign long-term contracts, in which case the results could also reflect adverse selection.

Of course, one way for a team to reduce the moral hazard response is to reward

---

14 Labor supply effects also appear in men’s professional tennis. The Association of Tennis Professionals (the governing body of the men’s pro tour) decided several years ago to consider just a player’s results in the most recent 14 events in which they have competed in computing his ranking, which affects his seeding and therefore success probability in future tournaments. For example, in 1999, one player was ranked first in the world for part of the year, despite a string of first round tournament losses. These did not count in his ranking point total, and observers surmised that he would have entered fewer tournaments (and thus conserved his energy) if the costs of losing were higher. Because of these poor incentives, the Association of Tennis Professionals (the ruling body of men’s tennis) has instituted a new system which will count in a player’s ranking a core set of tournaments, whether the player plays in the tournament or not, and then some other tournaments as well. See the ATP’s website at (http://www.atptour.com).
players for not being injured. Lehn (1990) notes that 38 out of 155 players with contracts of three or more years, or about 25 percent, had incentive clauses in their contracts. These clauses sometimes rewarded either being available to play for most of the season or postseason awards won (such awards typically require being active for all or most of the year). Before signing such long-term contracts, those who ended up with incentive bonuses had virtually identical average propensities to be injured as those without such incentive bonuses. However, after signing, the injured time of players without incentive bonuses was 2.4 times that of those with bonuses. Again, a strong moral hazard response is suggested, although as before, we cannot rule out the adverse selection possibility that players who suspected that they were likely to be fragile have turned down the opportunity to sign a contract with an incentive bonus.

Hiring better quality management is an alternative route, along with contract incentives, for eliciting desired performance levels. In a study of the impact of baseball managers, I estimated the effect of better managers on team and individual player performance (Kahn, 1993b). Managerial quality was measured by first running a 1987 regression with manager salary as the dependent variable and managerial experience, career winning percentage, and a National League dummy variable as the explanatory variables. Then, using the coefficients from the regression, I plugged in each manager’s actual experience and winning percentage to get a predicted salary level. I then calculated that during the 1969–86 period, hiring a better quality manager significantly raised the team’s winning percentage relative to its past level—even if one also controls for team scoring and runs allowed, suggesting that good managers win the close games. The effect of good managers was even larger when I didn’t control for offense and defense. The latter effect could indicate that better managers are superior judges of talent, or motivate their players, and thus indirectly contribute to offense and defense.

I also studied individual player performance relative to established career levels when the team was taken over by a new manager. The better the quality of the new manager, the better a player’s future performance relative to his past performance. In related calculations, I found an increase in managerial quality more than pays for itself, based on Scully’s (1989) results for the effect of winning on revenue. Because of this, one might have expected the salaries of highly talented managers to be bid up. The fact that they weren’t as measured in the 1987 salary data used in this study may be further indirect evidence of collusion between baseball owners during this time period.

**Final Thoughts**

Labor issues in sports may seem distant from the rest of the economy, since they often seem to pit millionaire players against billionaire owners. But while it would be unwise to extrapolate too strongly from the labor market experience of sports, evidence on a particular labor market should not be discounted just because the market has a high profile, either. The strong evidence for monopsony in sports has some parallels to a similar effect that has been found among groups such as public school teachers, nurses and university professors (Ehrenberg and Smith,
The evidence from these areas suggests that the phenomenon of employer monopsony power could be more widespread than is commonly acknowledged by economists. The presence of customer discrimination in sports reminds us that there are many sectors in the economy with producer-customer contact where discrimination could persist. The results on player performance suggest that athletes are motivated by similar forces that affect workers in general.

While this paper has concentrated on sports in North America, many of the same economic issues arise in the sports industry elsewhere. Professional soccer leagues in Europe are tremendously lucrative and also must be concerned with player movement and competitive balance. For example, television contracts for soccer in the United Kingdom, France, Italy, Spain and Germany had a combined annual value of roughly $1 billion in 1997; this compares favorably to the 1996 total media revenues of roughly $1.3 billion for the most lucrative league in North America, the NFL. In fact, European soccer draws more TV revenue than the NBA, Major League Baseball or the NHL (Hoehn and Szymanski 1999; Financial World, June 17, 1997). The promotion and demotion of individual teams to and from a new European superleague involving teams from several countries raise fascinating questions about the role of competitive balance (Hoehn and Szymanski, 1999).

Recent developments in North American sports will provide some additional opportunities to observe economic theories at work. For example, the NBA and the NFL have adopted explicit revenue sharing rules between labor and management so that salaries are supposed to comprise a certain share of total revenue. Major League Baseball has adopted revenue sharing and a “luxury tax” on payrolls over a particular amount. While explicit revenue sharing can promote cooperation between labor and management by automatically joining each side’s economic success together, the 1998 basketball lockout suggests that agreeing on an appropriate sharing rule can be difficult.

References


Becker, Brian E. and Mark A. Huselid. 1992. “The Incentive Effects of Tournament Compen-

The author thanks Robert F. Burk for supplying baseball salary data, Rodney Fort for attendance data, Ronald Ehrenberg, Alan Krueger, and Timothy Taylor for helpful comments and suggestions, and Abhijay Prakash for excellent research assistance. Portions of this paper were written while the author was a Visiting Scholar at the Russell Sage Foundation, to which he is grateful for research support.


